New insights from genome-wide association analysis using imputed whole-genome sequence: the genetic mechanisms underlying residual feed intake in chickens

Shaopan Ye

South China Agricultural University

Zitao Chen

South China Agricultural University

Rongrong Zheng

South China Agricultural University

Shuqi Diao

South China Agricultural University

Jinyan Teng

South China Agricultural University

Xiaolong Yuan

South China Agricultural University

Hao Zhang

South China Agricultural University

Zanmou Chen

South China Agricultural University

Xiquan Zhang

South China Agricultural University

Jiaqi Li

South China Agricultural University

Zhe Zhang (✉ zhezhang@scau.edu.cn)

South China Agricultural University <https://orcid.org/0000-0001-7338-7718>

Research article

Keywords: Whole genome sequence, GWAS, transcriptome analysis, feed efficiency, chickens.

Posted Date: October 1st, 2019

DOI: <https://doi.org/10.21203/rs.2.15454/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Poultry feed occupies the largest cost of poultry production, which is estimated up to 70%. Moreover, it is pressure on the agricultural industry to reduce emissions and improve its environmental footprint, simultaneously increasing output to meet the growing demand for protein worldwide. Therefore, improving feed efficiency (FE) play an important role to improve profits and their environmental footprint in broiler production. In this study, using imputed whole genome sequencing (WGS) data, genome-wide association analysis (GWAS) was performed to identify SNPs and genes associated with residual feed intake (RFI) and its component traits. Furthermore, transcriptomic analysis between high- and low-RFI groups was performed to validate candidate genes from GWAS. Results: Results showed that heritability estimates of average daily gain (ADG), average daily feed intake (ADFI) , and RFI were 0.29 (0.004), 0.37 (0.005), and 0.38 (0.004), respectively. Using imputed sequence-based GWAS, we identified seven significant SNPs and five candidate genes (MTSS I-BAR domain containing 1, MTSS1; folliculin, FLCN; COP9 signalosome subunit 3, COPS3; 5',3'-nucleotidase, mitochondrial, NT5M; and gametocyte specific factor 1, GTSF1) associated with RFI, twenty significant SNPs and one candidate gene (inositol polyphosphate multikinase, IPMK) associated with ADG, and one significant SNP and one candidate gene (coatomer protein complex subunit alpha, COPA) associated with ADFI. After performing transcriptomic analysis between high- and low-RFI groups, both 38 up-regulated and 26 down-regulated genes were identified in high-RFI group. Furthermore, integrating regional conditional GWAS and transcriptome analysis, ras related dexamethasone induced 1 (RASD1) was the only one overlapped gene associated with RFI, which also suggested that the region (GGA14: 4767015 -4882318) is a new quantitative trait locus (QTL) associated with RFI. Conclusions: In conclusions, using imputed sequence-based GWAS is an efficient method to identify significant SNPs and candidate genes in chicken. Our results provide valuable insights into the genetic mechanisms of RFI and component traits, which would further improve the genetic gain of feed efficiency rapidly and cost-effectively in the context of marker-assisted breeding selection. Keywords: Whole genome sequence, GWAS, transcriptome analysis, feed efficiency, chickens.

Background

Poultry feed occupies the largest cost of poultry production, which is estimated up to 70% [1]. Moreover, there is pressure on the agricultural industry to reduce emissions and improve its environmental footprint, simultaneously increasing output to meet the growing demand for protein worldwide. Therefore, improving feed efficiency (FE) play an important role to improve profits and their environmental footprint in broiler production. Residual feed intake (RFI), a sensitive and accurate measure of feed efficiency, has been widely used in the genetic improvement of FE in livestock. RFI is defined as the difference between actual and expected feed intake given body weight and weight gain [2]. The RFI methodology separated broiler breeder energy efficiency into two components: systematic sources of variation related to in individual maintenance, and all other sources of variation in energy efficiency (the residual) [3]. Compared with the genetic improvement of feed conversion ratios (FCR), the improvement of RFI may have a simultaneous positive effect on productivity and feed efficiency [4]. In addition, RFI has a moderate heritability in broilers and responds to selection [4-6]. Although traditional selection for RFI has made substantial genetic gain, the genetic mechanisms still unclear [7]. Therefore, genetic dissection of RFI and its component traits (average daily gain, ADG and average daily feed intake, ADFI) would further improve the genetic gain of feed efficiency rapidly and cost-effectively.

Over the past decade, the genome-wide association analysis (GWAS) based on SNP chip data to identify the genetic mechanisms of FE has been widely implemented in livestock, especially in cattle [8-10] and pig [11-13]. These GWAS studies revealed many candidate genes and provided useful information for genomic breeding programs to select more efficient animals in livestock. However, the RFI related GWAS studies is still scarce in chicken. Yuan et al. [14] performed a GWAS using 600 K SNP array, identified a haplotype block on GGA27 harbored a significant SNP (rs315135692) associated with RFI. Also using GWAS with 600 K SNP array, our previous study showed that the effective SNPs related RFI were located in a 1-Mb region (16.3–17.3 Mb) of GGA12, but did not identified causal variant SNPs or genes associated with RFI [15]. The poor power of our previous GWAS study maybe due to the small population size and lower marker density. Nowadays, the rapid development of high-throughput sequencing technology has made it possible to perform GWAS with whole genome sequencing (WGS) data. Compared with SNP chip data, WGS data include whole genomic variants as well as causal mutations. Thus, GWAS combined with WGS data is expected to improve the power of test efficiency and identify the causal mutations of complex traits, and this expectation has confirmed in dairy cattle populations [16]. However, sequencing thousands of individuals of interest are still too expensive. Imputation from SNP panels to WGS data is an attractive and less expensive approach to obtain WGS data [17].

In this study, using a combined reference panel, genotype imputation was performed to obtain WGS data with two-step approach (from 55 to 600 K, and then imputed to WGS). Using imputed WGS data, GWAS was performed to identify SNPs and genes associated with RFI and its component traits. In addition, transcriptomic analysis was performed to identifies differentially expressed genes (DEGs) between high- and low-RFI groups to validate candidate genes of RFI from GWAS results, and give biological evidence to candidate genes. The aims of this study were to pinpoint associated loci and genes that contribute to the phenotypic variability in feed efficiency, and provide valuable insights into the genetic mechanisms of RFI.

Results

Basic descriptive statistics of analyzed traits and genetic parameter estimations

For improving the genetic improvement of feed efficiency in broiler chickens, three traits were considered to analyze, including ADG, ADFI, and RFI. Basic descriptive statistics of these traits were shown in Table 1. The Shapiro-Wilk statistic of these traits were closed to 1, and approximate p-values less than 0.054, which indicated that obey gaussian distribution. And the standard deviation (SD) of RFI, ADG, and ADFI were 8.83, 4.87, and 14.36, respectively. In additional, the coefficient of variation (CV) were 13.22% of ADFI and 16.82% of ADG. Therefore, these traits existed substantial phenotypic variation (Table 1). Using the AI-REML method, estimates of heritability, phenotypic correlation, and genetic correlation among ADG, ADFI, and RFI were calculated and showed in Table 2. Results show that heritability estimates of ADG, ADFI, and RFI were 0.29 (0.004), 0.37 (0.005), and 0.38 (0.004), respectively. Genetic parameter analyses shown that ADFI were both positively and highly interrelated with ADG and RFI. In addition, RFI is poorly correlated with ADG both genetic relationship and phenotypic correlation.

Imputed sequence-based GWAS for ADG

Using the univariate model, sequence-based GWAS were performed for ADG, and the Manhattan and Quantile-Quantile (QQ) plots of GWAS results of ADG were shown in Figure 1A. Both 20 significant SNPs (Table 3) and 24

suggestive significant SNPs (Table S1) associated with ADG were identified using the threshold of suggestive and significant P-values (4.98×10^{-6} , and 2.49×10^{-7}). These SNPs were located on these chromosome (GGA1, GGA3, GGA6, GGA14, GGA25, and GGA27). And all significant SNPs were in high linkage disequilibrium (Figure.S2) and located in a region that ranged from 535.4 to 538.9 kb on GGA6. After stepwise conditional analysis, the level of all significant or suggestive loci around the lead SNP decreased below the genome-wide suggestive threshold in the conditional GWAS (Figure S3). And all significant SNPs were located in intergenic region, the nearest gene was inositol polyphosphate multikinase (IPMK) (Table 3). Additionally, the genomic inflation factor of ADG was 1.024, which indicated that the results of GWAS of ADG was acceptable.

Imputed sequence-based GWAS for ADFI

Using the univariate model, sequence-based GWAS were performed for ADFI, and the Manhattan and Quantile-Quantile (QQ) plots of GWAS results of ADFI were shown in Figure 1B. Both one significant SNPs (Table 3) and 140 suggestive significant SNPs (Table S1) associated with ADFI were identified using the threshold of suggestive and significant P-values (4.98×10^{-6} , and 2.49×10^{-7}). These SNPs were located on these chromosome (GGA1, GGA3, GGA4, GGA14, and GGA25). And the most significant SNPs located at 1,147,989bp position of GGA25. And highly linkage disequilibrium between SNPs of significant regions were found on GGA25 (Figure S4). After stepwise conditional analysis, the level of all significant or suggestive loci around the lead SNP (rs15997392) decreased below the genome-wide suggestive threshold in the conditional GWAS (Figure S5). A total of 7 genes were found in the region, which extended 50 kb flanking regions in both upstream and downstream of lead SNP (rs15997392) position (Figure S5). And the independent significant SNP (rs15997392) was an intron variant in coatomer protein complex subunit alpha (COPA) (Table 3). Additionally, the genomic inflation factor of ADFI was 1.024, which indicated that the results of GWAS was acceptable.

Imputed sequence-based GWAS for RFI

Using the univariate model, sequence-based GWAS was performed for RFI, and the Manhattan and Quantile-Quantile (QQ) plots of GWAS results of RFI were shown in Figure 1C. Both 7 genome-wide significant SNPs (Table 3) and 100 suggestively significant SNPs (Table S1) associated with RFI were identified using the threshold of suggestive and significant P-values (4.98×10^{-6} , and 2.49×10^{-7}). These significant SNPs were mainly located on these chromosome (GGA2, GGA14, and GGA27). And the most significant SNP was located at 4,779,635 bp on GGA14, which explained 5.46% of the phenotypic variance of RFI. Using SnpEff software, five candidate genes associated with RFI were identified and annotated (Table 3). There are MTSS I-BAR domain containing 1 (MTSS1), folliculin (FLCN), COP9 signalosome subunit 3 (COPS3), 5',3'-nucleotidase, mitochondrial (NT5M), and gametocyte specific factor 1 (GTSF1). After LD analysis, the highly linkage disequilibrium between significant SNPs was found on GGA14 and GGA27 (Figure S6, S7). To find the independent SNPs in a chromosome, stepwise conditional analysis was performed by adding the lead SNP to the model as fixed effect. The level of all significant or suggestive loci around the lead SNP decreased below the genome-wide suggestive threshold in the conditional GWAS (Figure 2A). Extended 50 kb flanking regions in both upstream and downstream of lead SNP position of GWAS results, five genes (ENSGALG00000004816, COP93, NT5M, MED9 (mediator complex subunit 9), ras related dexamethasone induced 1 (RASD1)) were indicated on GGA14 (Figure 2A), and six genes (ENSGALG00000027009, ENSGALG00000045610, ENSGALG00000027214, GTSF1, golgi SNAP receptor complex member 2 (GOSR2), and ENSGALG00000037637) on GGA27 (Figure 2B). In addition, the significant SNP of GGA2 was an isolated signal, which suggested it maybe a false positive significant site.

Therefore, only two independent SNPs (rs314690911 and rs317155749) were suggested significance association with RFI in this chicken population. Both substitutions variants of rs314690911 (A to G) and rs317155749 (G to A) led to a significant decrease of in RFI phenotypic value (Figure 3). Additionally, the genomic inflation factor of RFI was 1.002, which closes to 1.00 reflects the results of GWAS was acceptable.

Validation of candidate genes from GWAS results based on differentially expressed genes (DEGs)

To validate the list of candidate genes from GWAS results, we performed transcriptomic analysis to identify DEGs between high- and low-RFI groups in chicken. There were 64 genes differentially expressed between high- and low-RFI with gene expression fold change ranging from -5.33 to 5.20. Compared with low-RFI group, both 38 up-regulated genes and 26 down-regulated were identified in high-RFI group. All of significant DEGs between high- and low-RFI groups was shown in Figure.4 and Table S2, and the top 10 differentially expressed genes were monoamine oxidase A (MAOA), heat shock protein 90 beta family member 1 (HSP90B1), cytochrome P450 family 2 subfamily C polypeptide 23a (CYP2C23a), liver enriched antimicrobial peptide 2 (LEAP2), RASD1, ABI family member 3 binding protein (ABI3BP), and ADAMTS like 5 (ADAMTSL5). After functional and pathway enrichment analysis, two molecular function (carbon-carbon lyase activity and unfolded protein binding) and one cellular component (endoplasmic reticulum lumen) categories were significantly enriched (Table S3). Compared DEGs from transcriptomic analysis (Figure 4) with these candidate genes that located on expanding 50 kb flanking regions in both upstream and downstream of the independent SNPs rs314690911 and rs317155749 (Figure 2), RASD1 was the only one overlapped gene, which suggested a new QTL (GGA14:4767015-4882318) truly associated with RFI (Figure S1).

Comparison between significant SNPs from GWAS and reported QTLs

SNPs less than the threshold of genome-wide significant P-values (4.98×10^{-6}) were selected to compare with the reported QTLs. These QTLs were collected from the Animal QTL database based on their physical locations. For RFI, only two QTLs located on GGA4 and GGAZ were extracted, and the nearest distance of QTL between significant SNPs was 24,435,812 bp on GGAZ (Table S4). For ADG, a total of 4 QTLs located on autosomes (GGA 1, 3, 6 and 27) were extracted, and the nearest distance of QTL between significant SNPs was 2,588,827 bp on GGA 27 (Table S4). For ADFI, a total of 6 QTLs located on autosomes (GGA 1, 3 and 4) were extracted, and the nearest distance of QTL between significant SNPs was 1,504,353 bp on GGA 4 (Table S4). None significant SNPs located inside of reported QTL were found.

Discussion

Feed efficiency (FE) play an important role to improve profits and their environmental footprint in broiler production. The genetic dissection of RFI and its component traits would further improve the genetic gain of feed efficiency rapidly and cost-effectively. In this study, a chicken population with imputed WGS data was used to perform GWAS for exploring the genetic mechanisms of feed efficiency. To the best of our knowledge, this is the first time to perform GWAS for RFI in chicken using imputed sequence data, and performed transcriptomic analysis to identify DEGs between high- and low-RFI groups in chicken to validate these candidate genes.

GWAS with imputed whole genome sequence data

Recently, GWAS with imputed WGS data have been widely used in livestock species, especially in cattle [16, 18, 19]. Because of genotype imputation would improve the power of GWAS [20], and reduce the cost of genotyping. However, genotype imputation from SNP array to WGS data not only increased the marker density, but also brought the imputation error. And imputation errors will affect the probability of causal SNPs, which were determined by performing GWAS. Therefore, it is necessary to ensure the highly imputation accuracy before performing GWAS. For obtained higher imputation accuracy, genotype imputation with two-step approach was performed using a combined reference panel in this study. After quality control of the imputed WGS data, the average imputation accuracy (Beagle R²) of all SNPs was 0.871 ± 0.177, and ranged from 0.357 to 0.926 for different chromosomes (Table S5). This high imputation accuracy was full enough to ensure the confidence of GWAS results. This obtained high imputation accuracy may due to performed imputation with two-step approach [21] and using a combined reference panel [22]. Compared with our previous study, which performed GWAS for RFI using 426 chickens with 600 K array data [15], new significant SNPs and genes were identified. Because of more individuals (626 birds) and higher marker density (imputed WGS data) were utilized to perform GWAS. Although the power of GWAS was improved in this study, it is still difficult to pinpoint the causative variant. Because many significant SNPs existed a high degree of linkage disequilibrium, and had almost equally significantly associated variations, such as the significant SNPs of ADG on GGA6 (Table 3).

Candidate genes for residual feed intake (RFI) and its component traits

After performed sequence-based GWAS, a lot of candidate genes were annotated that associated with RFI and its component traits (Table 3 and Table S2). For RFI, two independent significance SNPs (rs314690911 and SNP rs317155749) were identified via conditional GWAS. One was a variant upstream of COPS3, and the other is an intron variant of GTSF1. The COPS3 gene encodes the third subunit of the eight-subunit COP9 signalosome complex, which was initially identified as a negative regulator of constitutive photomorphogenesis in *Arabidopsis thaliana* [23]. Also, COPS3 plays an important role in both the tumorigenesis and progression. Previous research shown that the amplification of COPS3 was strongly associated with large tumor size ($P = 0.0009$) [24]. The other is GTSF1, which encodes a 167-amino acid protein as a member of the Uncharacterized Protein Family 0224 (UPF0224), and play an important role in liver cancer. Compared with GTSF1-positive group, the sizes and weights of the tumors of liver cancer in the GTSF1-negative group were decreased significantly ($P < 0.05$) [25]. Moreover, RASD1 was the only one overlapped gene between GWAS results and DEGs between high- and low-RFI groups. RASD1 is a highly conserved member of the Ras family of monomeric G proteins that was first identified as a dexamethasone (DEX) inducible gene in AtT-20 mouse pituitary tumor cells [26]. Vaidyanathan et al. indicate that RASD1 often promotes cell growth [27]. Moreover, the current literature demonstrates that Rasd1 expression is induced by a wide range of physiological stimuli and has many biological effects including the regulation of circadian timekeeping, anxiety related behavior, adipocyte differentiation, and hormone release [28]. For ADG, all significant SNPs were located in intergenic region, and IPMK was the nearest gene (Table 3). IPMK is a member of the IPK-superfamily of kinases, which plays an important role at the nexus of signaling, metabolic and regulatory pathways [29]. For example, IPMK involved in hypothalamic control of food intake via AMP-activated protein kinase (AMPK) signaling pathways [30]. For ADFI, the most significant SNPs located at 1147989bp position of GGA25, which was an intron variant in COPA. COPA encodes the α -COP subunit of the coat protein I (COPI) seven subunit complex that is involved with intracellular coated vesicle transport [31]. COPA mutations have recently been revealed to cause autoimmune interstitial lung, joint and kidney disease (COPA syndrome).

Moreover, the COP9 signalosome is a subunit of a highly conserved complex of COPS3. Therefore, we guess the interaction of COPS3 and COP9 would impact the feed intake and further on RFI.

Combined GWAS and transcriptomic analysis to identify candidate genes

Nowadays, it is an efficient method that combined GWAS and transcriptomic analysis to identify the causal mutations of complex traits in livestock. In cattle, previous study integrated RNA-Seq data and sequence-based GWAS data to explore the genetic mechanisms of mastitis resistance and milk production [32]. In swine, combining the GWAS and gene expression profile data, two genes (UBA domain containing 1 gene and Epsin 1 gene) were identified significantly associated with *Streptococcus Suis* serotype 2 (SS2) resistance [33]. In chicken, integrated GWAS and transcriptome analysis, new finding about the molecular mechanism of chicken white/red earlobe color formation were revealed [34]. In this study, combining imputed sequence-based GWAS and transcriptome analysis between high- and low-RFI groups, we also found an overlapped gene (RASD1) significantly associated with RFI (Figure 3 and Figure 5). This finding also suggested that imputed sequence-based GWAS was an efficient method to identify significance SNPs.

Comparison between GWAS results and reported QTLs

Compared GWAS results with reported QTLs, we found that none significant SNPs located inside of reported QTL in this study (Table S4). This mainly due to the number of reported QTLs, which associated RFI, ADFI, and ADG, are still very limited in QTLdb, especial for RFI. In QTLdb, only 40 QTLs were reported associated with RFI, and 17 QTLs located in a range from 16433894 to 17287377 bp on GGA12. Moreover, the difference genetic backgrounds also would result the difference QTLs region.

Conclusions

In conclusions, using imputed sequence-based GWAS is an efficient method to identify significant SNPs and candidate genes in chicken. In this study, using imputed sequence-based GWAS, we identified seven significant SNPs associated with RFI, twenty significant SNPs associated with ADG, and one significant SNPs associated with ADFI. Furthermore, combined regional conditional GWAS and transcriptome analysis between high- and low-RFI groups, an overlapped gene (RASD1) were identified associated with RFI, which also suggested that a new QTL (GGA14:4767015-4882318) truly associated with RFI. Our results provide valuable insights into the genetic mechanisms of RFI and component traits in chickens.

Methods

Ethics statement

Animal care and experiments were conducted according to the Regulations for the Administration of Affairs Concerning Experimental Animals (Ministry of Science and Technology, China) and approved by the Animal Care and Use Committee of the South China Agricultural University, Guangzhou, China (approval number: SCAU#2014-10).

Population and phenotyping

A chicken population derived from a yellow-feather dwarf broiler breed that maintained for 25 generations by Wens Nanfang Poultry Breeding Co. Ltd. (Xinxing, P.R. China) was used in this study. This population includes 1,600 birds (800 males, 800 females) and was the 3rd batch of the 25th generation of this chicken population. These birds came from a mixture of full sib and half sib families from mating 30 males and 360 females of the 24th generation. After hatching, all birds were maintained in a closed building under controlled environmental conditions and provided with a standard diet until they were 4 wk of age. The chickens were then randomly assigned to six pens by gender (three pens for males and three pens for females) for growth performance testing from 5 to 13 wk of age. They received food and water *ad libitum* during all stages. Finally, remaining 1,338 birds were slaughtered at 91 day of age for carcass trait recording. Average daily gain (ADG) and average daily feed intake (ADFI) per individual were calculated for the period from 45 to 84 day. The and residual feed intake (RFI) were calculated by follow formula $RFI = ADFI - (b_0 + b_1 \times MMBW + b_2 \times ADG)$, where b_0 was the intercept, MMBW is mid-test body weight (MBW raised to the power of 0.75), and the MBW was the predicted body weight on day 21 of the test. b_1 and b_2 are the partial regression coefficients for MMBW and ADG, respectively. Descriptive statistics of analyzed traits could be found (Table 1). For more details about this population, please refer to Xu et al. (2016) and Zhang et al. (2017).

Genotyping, genotype imputation and quality control

After traits recorded systematically, a total of 644 male birds were randomly selected for genotyping. These birds were 15 male parents and 629 male offspring. Of these 644 birds, 450 birds were genotyped with the 600 K Affymetrix® Axiom® HD genotyping array [35], and remaining 194 birds were genotyped with the Affy 55K array [36]. The 600 K SNP chip contained 580,961 SNPs probes across 28 autosomes, two linkage groups (LGE64 and LGE22C19W28_E50C23), and two sex chromosomes. The 55 K SNP chip contained 52,184 SNPs probes across 28 autosomes and a sex chromosome (chrZ). After converting genome coordinates to a chicken reference genome (galGal5), 28 autosomes and a sex chromosome (chrZ) were extracted for further analyses. In the process of quality control of genotype, SNPs that minor allele frequency (MAF) smaller than 0.5%, genotyping call rate smaller than 97%, and Hardy–Weinberg equilibrium test P-value smaller than 1×10^{-6} were removed. Finally, 547,020 and 51,984 SNPs were remaining for 600 K and 55K chip data, respectively. In addition, a total of 23,213 SNPs was shared between 600 K and 55 K SNP chip.

Genotype imputation was performed with two-step approach from 55 to 600 K, and then imputed to WGS. Before Genotype imputation, pre-phasing was executed in Beagle 4.1 with default parameter [37]. Firstly, using 450 birds with 600 K chip data as a reference panel, these 194 birds were imputed from 55 K to 600 K chip data using Beagle 4.0 with pedigree. And then merge 600 K chip data of these 194 birds and 450 birds using VCFtools. Secondly, all of 644 birds with 600 K chip data were imputed to WGS data using a combined reference panel Beagle 4.1 with default parameter. The combined reference panels included 24 key individuals from the yellow-feather dwarf broiler population and 311 birds with WGS data from diverse chicken breeds. These 24 key individuals were selected by maximizing the expected genetic relationships [38]. These 311 birds were downloaded from 10 BioProjects in ENA or NCBI. The combined reference panels contained 36,840,795 SNPs probes across 28 autosomes and a sex chromosome (chrZ). More detail information could be found in our previous study [22].

After performed genotype imputation, quality control of the imputed WGS data was conducted using PLINK (Purcell et al., 2007) with the criteria of SNP call rate > 95%, individual call rate > 97%, MAF > 0.5%, and Hardy-

Weinberg equilibrium P-value > 1.0e-6. In addition, individuals would be excluded who existed Mendelian errors. Finally, the remaining 626 individuals and 11,173,020 SNPs were used for further analysis.

Genetic parameter estimations

The genomic heritability was calculated using the average information restricted maximum likelihood (AI-REML) method implemented in the software DMU v6.0 (Madsen and Jensen, 2013). The statistical model was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zg} + \mathbf{e}$$

where $\mathbf{y}$ is a vector of phenotypic values of all individuals; $\mathbf{b}$ is the vector of fixed effects including batch effect; $\mathbf{g}$ was the vector of the animal additive genetic effect, and $\mathbf{e}$ is the residual term, and $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices relating the fixed effects and the additive genetic values to the phenotypic records; $\mathbf{G}$ is the genomic relationship matrix calculated using 600 K chip data as follows [39]:

$$\mathbf{G} = \frac{\mathbf{MM}^T}{2 \sum_{i=1}^m p_i(1-p_i)}$$

where $\mathbf{M}$ is a matrix of centered genotypes, m is the number of markers, and p_i is the minor allele frequency of SNP i .

Genome-wide association analyses using imputed WGS data

Before performed GWAS, the population structure of this chicken population was calculated by PLINK. And found a slight population stratification (Figure S1), so that we added top five principal components as covariates into the GWAS model to adjust population structure. The univariate tests of association were performed using a mixed model approach implemented in the GEMMA v0.98.1 software [68]. All sequence variants after quality control were tested for associations. The model was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zg} + \mathbf{Sa} + \mathbf{e}$$

where $\mathbf{y}$ is a vector of phenotypic values of all individuals; $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices relating the fixed effects and the additive genetic values to the phenotypic records; $\mathbf{b}$ is the vector of fixed effects including batch effect and top five PCs; $\mathbf{g}$ is a vector of the genomic breeding values of all individuals; $\mathbf{a}$ is the additive effect of the candidate variants to be tested for association; $\mathbf{S}$ is a vector of the variants' genotype indicator variable coded as 0, 1, or 2; and $\mathbf{e}$ is the residual term. Genomic breeding values were assumed to be distributed as $\mathbf{g} \sim N(\mathbf{0}, \mathbf{G}\sigma_g^2)$, where $\mathbf{G}$ is the standardized relatedness matrix calculated by GEMMA using 600 K chip data. The Wald test was applied to test the alternative hypotheses of each SNP in the univariate models. And the variance contribution to additive genomic variance by a SNP was calculated as follows: $\sigma_{a_i}^2 = \frac{\beta_i^2}{2p_i(1-p_i)}$, where σ_a^2 is the additive genomic variance explained by a SNP, p_i is the allele frequency, and β_i is the SNP effect estimated from GWAS results.

The Manhattan plot and quantile-quantile plot (QQ plot) were generated by the qqman package (Turner 2014) in R. The threshold of genome-wide significant P-values was adjusted based on the effective number of independent tests for Bonferroni method. The imputed WGS data was pruned to 1,082,126 independent SNPs

using PLINK with the command (`-indep-pairwise 25 5 0.2`). The effective number of independent tests were set to 200,629, estimated by sampleM. Therefore, the genome-wide suggestive and significant P-values were 4.98×10^{-6} ($1.00/200,629$) and 2.49×10^{-7} ($0.05/200,629$), respectively. For evaluating the extent of false positive signals of GWAS results, a genomic inflation factor (λ) was calculated as the median of the resulting chi-squared test statistics divided by the expected median of the chi-squared distribution with one degree of freedom (i.e., 0.454). Haploview software (Barrett et al. 2005) was used to analyze the linkage disequilibrium around the significant SNPs. For identifying the independent signals precisely, the most significant SNP were added as covariates into the univariate models in step-wise conditional analyses. In addition, the gene information file of chicken was downloaded from Ensembl gene build 94, and candidate genes were annotated using the software SnpEff version 4.3t [40].

Transcriptomic analysis identifies differentially expressed genes (DEGs) associated with RFI

Raw reads of four samples (sample45561 and sample46307 with high RFI, and sample45012 and sample46732 with low RFI) were downloaded from our previous study [15]. And raw reads were processed to clean reads by filtering low quality reads and adaptor dimers. Clean reads were mapped to the chicken reference genome (galGal5) using HISAT2 v2.0.5 with the default parameter. Then, the alignments were assembled into full and partial transcripts using StringTie, and quantify the transcripts for each sample using the GAL5. Finally, the differential gene expression analysis was made with Ballgown in R environment [41]. In this study, differentially expressed transcripts or genes were identified based on an adjusted p-value less than 0.05 (false discovery rate, FDR of 5 %) and the absolute value of log2-transformed of fold change (FC) more than or equal to one. Function and pathway enrichment were performed by the R language package (clusterProfiler). Using Benjamini-Hochberg method, the P-values of KEGG pathway and GO terms were adjusted for multiple testing [42]. And an adjusted P-value less than 0.05 was set as significant P-values. In addition, the genome annotation information file was downloaded from Ensembl gene build 94.

Validation of candidate genes based on differentially expressed genes (DEGs)

The identification of candidate genes of lead SNPs from GWAS results was performed basing on their corresponding genomic positions. The candidate gene regions were defined as extended 50 kb flanking regions in both upstream and downstream of lead SNP position. If there are no genes in candidate gene regions, selected the nearest genes in both upstream and downstream of lead SNP position as the candidate genes. Compared the overlap genes or regions between the candidate genes from sequence-based GWAS and DEGs, to identify the causal genes or QTLs.

Significant SNPs compared with reported QTLs

To compare results from sequence-based GWAS with reported QTLs, significant and suggestively significant SNPs were selected to compare with the QTLs. These QTLs, all of which affect ADG, ADFI and RFI, were selected from the Animal QTLdb [43], respectively. QTLs closest to the significant SNPs were extracted.

Abbreviations

ABI3BP: ABI family member 3 binding protein; ADAMTSL5: ADAMTS like 5; ADG: average daily gain; ADFI: average daily feed intake; chrZ: a sex chromosome; COPS3: COP9 signalosome subunit 3; COPA: coatomer

protein complex subunit alpha; CYP2C23a: cytochrome P450 family 2 subfamily C polypeptide 23a; DEGs: identifies differentially expressed genes; ENA: European nucleotide archives; FDR: false discovery rate; FC: fold change; FE: feed efficiency; FLCN: folliculin; G: genomic relationship matrix; GTSF1: gametocyte specific factor 1; GWAS: genome-wide association analysis; GOSR2: golgi SNAP receptor complex member 2; HSP90B1: heat shock protein 90 beta family member 1; IPMK inositol polyphosphate multikinase; LEAP2: liver enriched antimicrobial peptide 2; QQ plot: quantile-quantile plot; QTL: quantitative trait locus; MMBW: mid-test body weight; MED9: mediator complex subunit 9; MAOA: monoamine oxidase A; MTSS1: MTSS I-BAR domain containing 1; MAF: minor allele frequency; NT5M: 5',3'-nucleotidase, mitochondrial; NCBI: national center for biotechnology information; RFI: residual feed intake; RASD1: ras related dexamethasone induced 1; WGS: whole genome sequencing; UPF0224: Uncharacterized Protein Family 0224.

Declarations

Consent for publication

Not applicable.

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (31772556), the earmarked fund for China Agriculture Research System (CARS-35, CARS-41).

Authors' contributions

SPY, ZZ, and JQL conceived the study and designed the project and helped draft. QXZ provided the chickens' dataset. SPY and RRZ finished genotype imputation and performed GWAS. SPY and ZTC performed transcriptomic analysis. ZTC, RRZ, SQD, JYT, YXL, HZ, and ZMC participated in the design, and contributed to the manuscript. All authors read and approved the manuscript.

Acknowledgements

We thank Zhengqiang Xu for providing RNA-sequencing data.

References

1. Willems OW, Miller SP, Wood BJ: **Aspects of selection for feed efficiency in meat producing poultry.** *World's Poultry Science Journal* 2013, **69**(01):77-88.
2. Chambers D, Gregory KE, Swiger LA, Koch RM: **Efficiency of Feed Use in Beef Cattle.** *Journal of animal science* 1963, **22**(2):486-494.

3. Romero LF, Zuidhof MJ, Renema RA, Robinson FE, Naeima A: **Nonlinear mixed models to study metabolizable energy utilization in broiler breeder hens.** *Poultry Sci* 2009, **88**(6):1310-1320.
4. Aggrey SE, Karnuah AB, Sebastian B, Anthony NB: **Genetic properties of feed efficiency parameters in meat-type chickens.** *Genetics Selection Evolution* 2010, **42**.
5. Pakdel A, Arendonk JAMV, Vereijken ALJ, Bovenhuis H: **Genetic parameters of ascites-related traits in broilers: correlations with feed efficiency and carcass traits.** *British Poultry Science* 2005, **46**(1):43-53.
6. Zhang Z, Xu ZQ, Luo YY, Zhang HB, Gao N, He JL, Ji CL, Zhang DX, Li JQ, Zhang XQ: **Whole genomic prediction of growth and carcass traits in a Chinese quality chicken population.** *Journal of animal science* 2017, **95**(1):72-80.
7. Tallentire CW, Leinonen I, Kyriazakis I: **Breeding for efficiency in the broiler chicken: A review.** *Agronomy for Sustainable Development* 2016, **36**(4).
8. Seabury CM, Oldeschulte DL, Saatchi M, Beever JE, Decker JE, Halley YA, Bhattarai EK, Molaei M, Freetly HC, Hansen SL *et al*: **Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle.** *BMC Genomics* 2017, **18**(1):386.
9. Schweer KR, Kachman SD, Kuehn LA, Freetly HC, Pollak JE, Spangler ML: **Genome-wide association study for feed efficiency traits using SNP and haplotype models.** *Journal of animal science* 2018, **96**(6):2086-2098.
10. Higgins MG, Fitzsimons C, McClure MC, McKenna C, Conroy S, Kenny DA, McGee M, Waters SM, Morris DW: **GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle.** *Scientific Reports* 2018, **8**(1):14301.
11. Do DN, Ostensen T, Strathe AB, Mark T, Jensen J, Kadarmideen HN: **Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs.** *BMC genetics* 2014, **15**(1):27.
12. Bai C, Pan Y, Wang D, Cai F, Yan S, Zhao Z, Sun B: **Genome-wide association analysis of residual feed intake in Junmu No. 1 White pigs.** *Anim Genet* 2017, **48**(6):686-690.
13. Horodyska J, Hamill RM, Varley PF, Reyer H, Wimmers K: **Genome-wide association analysis and functional annotation of positional candidate genes for feed conversion efficiency and growth rate in pigs.** *PloS one* 2017, **12**(6):e0173482.
14. Yuan J, Wang K, Yi G, Ma M, Dou T, Sun C, Qu L-J, Shen M, Qu L, Yang N: **Genome-wide association studies for feed intake and efficiency in two laying periods of chickens.** *Genetics Selection Evolution* 2015, **47**.
15. Xu Z, Ji C, Zhang Y, Zhang Z, Nie Q, Xu J, Zhang D, Zhang X: **Combination analysis of genome-wide association and transcriptome sequencing of residual feed intake in quality chickens.** *BMC Genomics* 2016, **17**(1):594.
16. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brondum RF, Liao X, Djari A, Rodriguez SC, Grohs C *et al*: **Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle.** *Nature genetics* 2014, **46**(8):858-865.
17. Li Y, Willer C, Sanna S, Abecasis G: **Genotype imputation.** *Annual Review of Genomics & Human Genetics* 2009, **10**(10):387.
18. Iso-Touru T, Sahana G, Guldbbrandtsen B, Lund MS, Vilkkki J: **Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants.** *BMC genetics* 2016, **17**:55.

19. Littlejohn MD, Tiplady K, Fink TA, Lehnert K, Lopdell T, Johnson T, Couldrey C, Keehan M, Sherlock RG, Harland C *et al*: **Sequence-based Association Analysis Reveals an MGST1 eQTL with Pleiotropic Effects on Bovine Milk Composition.** *Sci Rep* 2016, **6**:25376.
20. Marchini J, Howie B: **Genotype imputation for genome-wide association studies.** *Nature Reviews Genetics* 2010, **11**:499.
21. Kreiner-Moller E, Medina-Gomez C, Uitterlinden AG, Rivadeneira F, Estrada K: **Improving accuracy of rare variant imputation with a two-step imputation approach.** *Eur J Hum Genet* 2015, **23**(3):395-400.
22. Ye S, Yuan X, Huang S, Zhang H, Chen Z, Li J, Zhang X, Zhang Z: **Comparison of genotype imputation strategies using a combined reference panel for chicken population.** *Animal : an international journal of animal bioscience* 2019, **13**(6):1119-1126.
23. Kwok SF, Solano R, Tsuge T, Chamovitz DA, Ecker JR, Matsui M, Deng XW: **Arabidopsis homologs of a c-Jun coactivator are present both in monomeric form and in the COP9 complex, and their abundance is differentially affected by the pleiotropic cop/det/fus mutations.** *The Plant cell* 1998, **10**(11):1779-1790.
24. Yan T, Wunder JS, Gokgoz N, Gill M, Eskandarian S, Parkes RK, Bull SB, Bell RS, Andrulis IL: **COPS3 amplification and clinical outcome in osteosarcoma.** *Cancer* 2007, **109**(9):1870-1876.
25. Gao DY, Ling Y, Lou XL, Wang YY, Liu LM: **GTSF1 gene may serve as a novel potential diagnostic biomarker for liver cancer.** *Oncology letters* 2018, **15**(3):3133-3140.
26. Kempainen RJ, Behrend EN: **Dexamethasone rapidly induces a novel ras superfamily member-related gene in AtT-20 cells.** *The Journal of biological chemistry* 1998, **273**(6):3129-3131.
27. Vaidyanathan G, Cismowski MJ, Wang G, Vincent TS, Brown KD, Lanier SM: **The Ras-related protein AGS1/RASD1 suppresses cell growth.** *Oncogene* 2004, **23**(34):5858-5863.
28. Bouchard-Cannon P, Cheng H-YM: **RASD1.** In: *Encyclopedia of Signaling Molecules*. Edited by Choi S. New York, NY: Springer New York; 2017: 1-9.
29. Kim E, Beon J, Lee S, Park J, Kim S: **IPMK: A versatile regulator of nuclear signaling events.** *Adv Biol Regul* 2016, **61**:25-32.
30. Lee J-Y, Kim Y-r, Park J, Kim S: **Inositol polyphosphate multikinase signaling in the regulation of metabolism.** *Annals of the New York Academy of Sciences* 2012, **1271**(1):68-74.
31. Watkin LB, Jessen B, Wiszniewski W, Vece TJ, Jan M, Sha Y, Thamsen M, Santos-Cortez RL, Lee K, Gambin T *et al*: **COPA mutations impair ER-Golgi transport and cause hereditary autoimmune-mediated lung disease and arthritis.** *Nature genetics* 2015, **47**(6):654-660.
32. Fang L, Sahana G, Su G, Yu Y, Zhang S, Lund MS, Sorensen P: **Integrating Sequence-based GWAS and RNA-Seq Provides Novel Insights into the Genetic Basis of Mastitis and Milk Production in Dairy Cattle.** *Sci Rep* 2017, **7**:45560.
33. Ma Z, Zhu H, Su Y, Meng Y, Lin H, He K, Fan H: **Screening of Streptococcus Suis serotype 2 resistance genes with GWAS and transcriptomic microarray analysis.** *BMC Genomics* 2018, **19**(1):907.
34. Luo W, Xu J, Li Z, Xu H, Lin S, Wang J, Ouyang H, Nie Q, Zhang X: **Genome-Wide Association Study and Transcriptome Analysis Provide New Insights into the White/Red Earlobe Color Formation in Chicken.** *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* 2018, **46**(5):1768-1778.

35. Kranis A, Gheyas AA, Boschiero C, Turner F, Yu L, Smith S, Talbot R, Pirani A, Brew F, Kaiser P *et al*: **Development of a high density 600K SNP genotyping array for chicken**. *BMC Genomics* 2013, **14**(1):59.
36. Liu R, Xing S, Wang J, Zheng M, Cui H, Crooijmans R, Li Q, Zhao G, Wen J: **A new chicken 55K SNP genotyping array**. *BMC Genomics* 2019, **20**(1):410.
37. Browning B, Browning S: **Genotype Imputation with Millions of Reference Samples**. *American journal of human genetics* 2016, **98**(1):116-126.
38. Ye SP, Yuan XL, Lin XR, Gao N, Luo YY, Chen ZM, Li JQ, Zhang XQ, Zhang Z: **Imputation from SNP chip to sequence: a case study in a Chinese indigenous chicken population**. *Journal of animal science and biotechnology* 2018, **9**.
39. VanRaden PM: **Efficient methods to compute genomic predictions**. *Journal of dairy science* 2008, **91**(11):4414-4423.
40. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM: **A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3**. *Fly* 2012, **6**(2):80-92.
41. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL: **Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown**. *Nature protocols* 2016, **11**(9):1650-1667.
42. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing**. *Journal of the Royal Statistical Society Series B (Methodological)* 1995, **57**(1):289-300.
43. Hu ZL, Fritz ER, Reecy JM: **AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond**. *Nucleic acids research* 2007, **35**(Database issue):D604-609.

Tables

Table 1. Basic descriptive statistics of analyzed traits

Trait ¹	N	Mean	SD	C.V.	Min	Max	W ²	P.value
ADG, g/d	626	28.95	4.87	16.82%	4.66	44.65	0.99	0.003
ADFI, g/d	626	108.59	14.36	13.22%	69.79	155.92	1.00	0.054
RFI, g	626	0	8.83	-	-32.05	26.00	0.99	0.003

¹These traits were average daily gain (ADG), average daily feed intake (ADFI), and residual feed intake (RFI).

²W means Shapiro-Wilk statistic.

Table 2. Genetic parameters of analyzed traits estimated by DMU with 600 K chip data in chicken

Trait ¹	ADG	ADFI	RFI
ADG	0.29 (0.004)	0.65	-0.01
ADFI	0.68 (0.09)	0.37 (0.005)	0.63
RFI	0.17 (0.16)	0.75 (0.07)	0.38 (0.004)

¹ADG: average daily gain; ADFI: average daily feed intake; RFI: residual feed intake. These diagonal values (mean $\pm$ S.E), lower triangles values (mean $\pm$ S.E), and upper triangles values are heritability estimates, genetic correlations, and phenotypic correlations for analyzed traits, respectively.

Table 3 The information of significant SNPs associated with ADG, ADFI and RFI

Trait*	SNPs	Chr.	Position	Allele	MAF	Bate (S.E)	- log10(P.value)		Candidate or nearest genes	Annotation
ADG	6:5354190	6	5354190	T/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs731666382	6	5363742	T/C	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs731606971	6	5366469	G/C	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs733679573	6	5366491	T/C	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs315729276	6	5366673	G/T	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs732899645	6	5367289	T/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5367311	6	5367311	T/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs315794025	6	5367445	G/T	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs739465741	6	5368822	G/C	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs731318403	6	5368918	A/T	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5376152	6	5376152	T/A	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5382202	6	5382202	T/C	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs731882620	6	5386041	G/T	0.009	-7.32(1.39)	0.96	6.74	IPMK	Intergenic
ADG	6:5386085	6	5386085	A/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5388772	6	5388772	C/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs736396454	6	5389064	A/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5389116	6	5389116	A/G	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	6:5389139	6	5389139	G/A	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs732888314	6	5389223	C/T	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADG	rs741516185	6	5389362	G/A	0.008	-8.1(1.44)	1.04	7.53	IPMK	Intergenic
ADFI	rs15997392	25	1147989	A/G	0.248	-5.42(1.01)	10.96	6.9	COPA	Intron
RFI	rs313664593	2	1.39E+08	A/G	0.014	-10.5(1.96)	3.04	6.93	MTSS1	Intron
RFI	rs313288641	14	4767015	G/A	0.107	-4.41(0.85)	3.72	6.61	FLCN	5 prime UTR
									COPS3	
										Upstream
RFI	rs314690911	14	4779635	A/G	0.142	-4.04(0.75)	3.98	6.96	COPS3	Upstream
									NT5M	Upstream
RFI	rs741733192	14	4782376	G/A	0.144	-3.98(0.75)	3.91	6.82	COPS3	Upstream
									NT5M	Upstream
RFI	rs314351418	14	4782740	G/A	0.144	-3.98(0.75)	3.91	6.82	COPS3	Upstream
									NT5M	Upstream
RFI	rs317155749	27	1212264	G/A	0.460	2.86(0.54)	4.06	6.85	GTSF1	intron
RFI	rs735238610	27	1220239	A/G	0.462	2.86(0.54)	4.07	6.81	GTSF1	intron

* ADG: average daily gain; ADFI: average daily feed intake; RFI: residual feed intake; is the additive genomic variance explained by a SNP

Figures

Figure 1

Manhattan plots and Q-Q plots of imputed sequence-based genome-wide association study for ADG, ADFI, and RFI. The Manhattan plots indicate $-\log_{10}(\text{observed P-values})$ for markers (y-axis) against their corresponding position on each chromosome (x-axis), while the Q-Q plots show the expected $-\log_{10}(\text{P-values})$ vs. the observed $-\log_{10}(\text{P-values})$. The horizontal blue and red lines represent the genome-wide significant threshold (4.98×10^{-6}) and genome-wide suggestive significant threshold (2.49×10^{-7}), respectively. Lambda represents genomic inflation factor.

Figure 2

Regional association plot of the lead SNP associated with RFI at GGA14 and GGA27. The regional association plot indicates $-\log_{10}(\text{observed P-values})$ for markers (y-axis) against their corresponding position on each chromosome (x-axis). The horizontal blue and red lines represent the genome-wide significant threshold (4.98×10^{-6}) and genome-wide suggestive significant threshold (2.49×10^{-7}), respectively. The lead SNPs are denoted by large black circle. SNPs are represented by colored circle according to the degree LD between the lead SNP. These green lines represent these genes located on this chromosome.

Figure 3

Genotype effect plot of lead SNP among three types at GGA14 and GGA27. (A) Genotype effect plot of lead SNP (rs314690911) among three types at GGA14. (B) Genotype effect plot of lead SNP (rs317155749) among three types at GGA27.

Figure 4

Differentially expressed genes (DEGs) between high- and low-RFI groups in chicken. The volcano plot indicates $-\log_{10}$ (observed P-values) for genes (y-axis) against their corresponding $\log_2$ (fold change) of echo gene (x-

axis). The horizontal red dotted line represents the significant threshold (0.05). The red, blue, and gray points represent up-regulated, down-regulated, and non-regulated genes in high-RFI groups, respectively.

Supplementary Files

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

- [Supplementaryfigure.docx](#)
- [Supplementarytables.xlsx](#)