A Genome-Wide SNP Scan Reveals Novel Loci for Egg Production and Quality Traits in White Leghorn and Brown-Egg Dwarf Layers

Wenbo Liu, Dongfeng Li, Jianfeng Liu, Sirui Chen, Lujiang Qu, Jiangxia Zheng, Guiyun Xu, Ning Yang

Abstract

Availability of the complete genome sequence as well as high-density SNP genotyping platforms allows genome-wide association studies (GWAS) in chickens. A high-density SNP array containing 57,636 markers was employed herein to identify associated variants underlying egg production and quality traits within two lines of chickens, i.e., White Leghorn and brown-egg dwarf layers. For each individual, age at first egg (AFE), first egg weight (FEW), and number of eggs (EN) from 21 to 56 weeks of age were recorded, and egg quality traits including egg weight (EW), eggshell weight (ESW), yolk weight (YW), eggshell thickness (EST), eggshell strength (ESS), albumen height (AH) and Haugh unit (HU) were measured at 40 and 60 weeks of age. A total of 385 White Leghorn females and 361 brown-egg dwarf dams were selected to be genotyped. The genome-wide scan revealed 8 SNPs showing genome-wise significant (P<1.51E-06, Bonferroni correction) association with egg production and quality traits under the Fisher’s combined probability method. Some significant SNPs are located in known genes including GRB14 and GALNT1 that can impact development and function of ovary, but more are located in genes with unclear functions in layers, and need to be studied further. Many chromosome-wise significant SNPs were also detected in this study and some of them are located in previously reported QTL regions. Most of loci detected in this study are novel and the follow-up replication studies may be needed to further confirm the functional significance for these newly identified SNPs.

Introduction

Modern egg layers have been selected for egg production and quality for decades. A breeding program for egg-type chickens has to face two major problems, the measurement of phenotypic data for individual hens and the efficient selection of cockerels which do not have records on laying performance and egg quality. Although great success has been achieved in layer breeding over the past decades, it is becoming more and more difficult with traditional breeding programs to make improvement in egg production and quality traits. With advances in technologies of molecular genetics and availability of DNA markers, identifying QTL controlling egg production and quality traits for application in marker-assisted breeding programs to make improvement in egg production and quality traits [4, 5, 6]. In addition, through making association test based on direct genotypic effects of markers within or near candidate genes, several polymorphisms were revealed relevant to egg production and quality traits [4, 5, 6].

Although many studies have successfully identified lots of QTL and a few causative genes, application of these results in commercial lines is still infeasible due to the precision of mapping. Identification of numerous single nucleotide polymorphisms in animal genome [7, 8], advances in high-throughput genotyping methods [9], and progresses in developing computational methods for analyzing high-density SNP data [10] have made possible using genomic information in livestock breeding. The successes of genome-wide association studies (GWAS) for detection of loci affecting milk production, fertility and growth traits in cattle [11, 12, 13] has spurred interest in the use of high-density SNP genotyping platform for the identification of sequence variations influencing egg production and quality traits in chicken. A 60 K SNP Illumina iSelect chicken array developed by the USDA Chicken GWMAS Consortium is a new and powerful platform for polymorphism detection in the whole genome of the chicken. In this study, we performed a GWA study to discover genomic regions explaining variations in egg production and quality traits of White Leghorn and brown-egg dwarf layers developed at China...
Agricultural University [6,14,15] by using the 60K SNP Illumina chicken array.

Materials and Methods

The blood samples were collected from brachial veins of chickens by standard venipuncture along with the regular quarantine inspection of the experimental station of China Agricultural University, and the whole procedure was carried out in strict accordance with the protocol approved by the Animal Welfare Committee of China Agricultural University (Permit Number: XK622).

Animals and Data Collection

Two experimental lines of egg-type chickens, White Leghorn (WL) and dwarf brown egg layer (DW), have been selected for egg production for over 10 years in the experimental station of China Agricultural University. These two lines were employed as two different experimental populations in the study. For each line, 600 hens from 40 families were kept in individual cages and their egg productions were recorded daily from 21 to 56 weeks of age. Production traits including total number of eggs from 21 to 56 wk of age (EN), age at first egg(AFE), first egg weight (FEW) were summarized. Individual data for egg quality traits including egg weight (EW), eggshell weight (ESW), Yolk weight (YW), eggshell thickness (EST), eggshell strength (ESS), albumen height(AH) and Haugh unit(HU) were measured with conventional methods. Eggs were collected in 3 consecutive days when hens were 40 weeks and 60 weeks old. The average for 3 d was taken as phenotypic value of each trait for every hen. As egg production decreased with age, some hens laid only one egg or even none within the 3 test days at 40 or 60 wks, and therefore no egg quality data were available for those layers at the respective age. All phenotypic values of traits in genotyped individuals were tested for normality, and some abnormal values extremely deviating from normal distribution were deleted. Box-Cox transformation was made for traits apart from normal distribution before conducting association tests.

Genotyping and Quality Control

Genomic DNA was isolated from blood sample by using standard phenol-chloroform extraction. Within family, two full-sib individuals with the same dam ID were randomly selected for SNP genotyping. In total, 385 White Leghorn females and 361 dwarf females were genotyped for 57,636 markers by using 60K SNP Illumina iSelect chicken array. These markers cover twenty-nine autosomes including GGA1 to 28 and GGA 32, two linkage groups containing E22C19W28_E50C23 and E64, and two sex chromosomes I and W are few in the 60K SNP array and most of them were deleted. Box-Cox transformation was made for traits apart from normal distribution before conducting association tests.

Statistical Methods for Association Study

Association tests were performed based on generalized least squares (GLS) testing to account for sib correlation by using EPISNP computer package [16], which is applicable to all bi-allelic loci of diploid species, so Z chromosome loci in the female individual of chicken cannot be analyzed. Moreover, GLS test can also adjust phenotypic observation of individuals for family structure before significance test for each SNP. According to the manual of EPISNP, the statistical tests followed a two-step least square analysis. For the first step, phenotypic values were corrected for fixed non-genetic effects, but in the present study all experimental chickens came from one generation with the same gender and they were raised in the same house under same environments, so there was no known fixed non-genetic effect impacting the results of association tests. In the second step, single-locus tests using the phenotypic values were conducted. The statistical model is:

\[ y = \mu + \text{SNP} + e \]

where \( y \) = phenotypic value corrected by fixed non-genetic effect, \( \mu \) = common mean, \( \text{SNP} \) = the single-locus SNP genotypic effect, and \( e \) = random residual. The single-locus SNP genotypic value was partitioned into additive and dominance effect. The extended Kempthorne model was applied for testing additive and dominant effects of each SNP in EPISNP [17]. In this model, a t-test was used to test the significance of additive and dominance using the following formula:

\[ t_s = \frac{L_s}{\sqrt{\text{var}(L_s)}} = \frac{s_g}{\sqrt{s_i(X'X)^{-1}s_i^2}} \]

where \( L_s = \text{contrast to estimate the genetic effect, } s_i = \text{a function of marginal and conditional allelic and genotypic frequencies for estimating additive or dominance effect, } s_g = \text{the least squares estimates of the SNP genotypic effects, } s^2 = (y - X\hat{g})' (y - X\hat{g})/(n - k) = \text{estimated residual variance, } y = \text{vector of phenotypic values, } X = \text{the design matrix, } n = \text{number of observations, } k = \text{rank of } X \ [16,17]. \]

White Leghorn and dwarf layers were analyzed independently. Results obtained from EPISNP were checked by using the linear model in PLINK [18]. As White Leghorn and dwarf layers are different in their genetic background and phenotypic values (Table 2), it is not suitable to incorporate data from these two populations in simultaneous association tests. A method for combined \( P \) values originally proposed by Fisher from independent tests of significance is adopted herein to address the heterogeneity of raw data. This strategy has also been successfully applied to GWAS and expression arrays elsewhere [19,20,21]. In this study, for a SNP marker, let \( P_{\text{ald}} \) and \( P_{\text{dw}} \) be the \( P \) value for significance of SNP effect (additive or dominance) obtained from association tests in White Leghorn and dwarf layer population separately, and \( s = -2\ln(P_{\text{ald}}+P_{\text{dw}}) \). Then under H0, \( P_{\text{aldw}} \sim \text{unif (0,1). Hence } s \sim \chi^2(df = 4) \). H0 can be rejected at \( \alpha \) level of significance if \( s > \chi^2(1-\alpha, df = 4) \). Bonferroni method was adopted to adjust for multiple testing from the number of SNP markers detected. A significant SNP was declared if its combined \( P \) value <0.05/N, here N is the number of SNP markers tested in combined analyses.

Results

Descriptive statistics of phenotypic measurements of egg quality and production traits in White Leghorn and dwarf layers used for GWAS were given in Table 2. All non-normal phenotypic data are within normal ranges after the Box-Cox transforming (Table 3).
Association tests were performed separately in White Leghorn and dwarf layer populations, and all results obtained from EPISNP were re-analyzed with PLINK. As PLINK also uses least squares regression analysis for quantitative traits, the results from PLINK were almost the same as those from EPISNP. Subsequently, P values from the two independent analyses for WL and DW were combined under the Fisher’s method by using 33,068 markers shared in the two populations. Taking 1.51E-06 (0.05/N, N = 33,068) as the genome-wise significance level with Bonferroni correction, it was revealed that 8 additive SNP effects showed significant association with egg production and quality traits including AFE, EN, ESW40, ESW60, EST40 and YW40. No dominance effects reached genome-wise significance. The profiles of P values, in terms of $-\log(P)$, of all tested SNPs after combining for different traits are shown in Figure 1. Furthermore, considering that Bonferroni correction is overly conservative and may lead to high proportion of negative false as marker density increase [22], we also tested association at chromosome-wide significance level, and the threshold ranged from 9.12E-06 on GGA1 to 6.67E-04 on E22 linkage group (Table S1). A total of 95 additive SNP effects

Table 1. Distributions of SNPs in 60k SNP Illumina iSelect chicken array and their conditions after quality control.

| GGA | No. SNP in chip | No. SNP remained after quality control | Average distance (kb) |
|-----|----------------|--------------------------------------|-----------------------|
|     |                | WL                                   | DW                    |
| 1   | 9059           | 6086                                 | 7163                  |
| 2   | 6958           | 4324                                 | 5387                  |
| 3   | 5171           | 3476                                 | 4072                  |
| 4   | 4256           | 2907                                 | 3418                  |
| 5   | 2766           | 1872                                 | 2171                  |
| 6   | 2219           | 1549                                 | 1710                  |
| 7   | 2280           | 1442                                 | 1742                  |
| 8   | 1813           | 1268                                 | 1341                  |
| 9   | 1504           | 1036                                 | 1182                  |
| 10  | 1682           | 1078                                 | 1275                  |
| 11  | 1647           | 929                                  | 1192                  |
| 12  | 1671           | 1111                                 | 1280                  |
| 13  | 1492           | 1043                                 | 1180                  |
| 14  | 1284           | 859                                  | 1023                  |
| 15  | 1337           | 855                                  | 1028                  |
| 16  | 31             | 18                                   | 22                    |
| 17  | 1104           | 730                                  | 891                   |
| 18  | 1160           | 742                                  | 837                   |
| 19  | 1053           | 709                                  | 792                   |
| 20  | 1991           | 1233                                 | 1375                  |
| 21  | 970            | 666                                  | 725                   |
| 22  | 476            | 231                                  | 293                   |
| 23  | 794            | 519                                  | 592                   |
| 24  | 937            | 645                                  | 736                   |
| 25  | 241            | 128                                  | 164                   |
| 26  | 852            | 586                                  | 660                   |
| 27  | 665            | 420                                  | 497                   |
| 28  | 793            | 487                                  | 563                   |
| 32  | 1              | 0                                    | 0                     |
| *E22 | 164           | 93                                   | 103                   |
| *E64 | 8             | 2                                    | 4                     |
| Z   | 3195           | 7                                    | 10                    |
| W   | 7              | 0                                    | 0                     |
| 10  | 1211           | 467                                  | 563                   |
| Total| 57636          | 37518                                | 43991                 |

1Conditions after quality control.
2White Leghorn.
3Dwarf brown egg layer.
4Linkage group.
5These SNPs are not assigned to any chromosomes.
doi:10.1371/journal.pone.0028600.t001
Table 2. Traits analyzed in White Leghorn and dwarf layers with phenotypic mean, standard deviation, and number of chickens with records.

| Trait                           | White Leghorn | Dwarf layer |
|---------------------------------|---------------|-------------|
|                                 | Mean  | SD    | n   | Mean | SD   | n   |
| Number of egg at 21–56 wk (EN)  | 196.0 | 26.0  | 385 | 194.2 | 23.5 | 361 |
| Age at first egg (AFE, day)     | 152.1 | 9.2   | 385 | 148.7 | 22.8 | 361 |
| First egg weight (FEW, g)       | 39.80 | 5.20  | 385 | 38.58 | 7.62 | 361 |
| Egg weight at 40 wk (EW40, g)   | 55.46 | 3.80  | 385 | 53.78 | 3.96 | 361 |
| Egg weight at 60 wk (EW60, g)   | 59.90 | 4.06  | 275 | 59.36 | 5.07 | 308 |
| Egg shell weight at 40 wk (ESW40,g) | 7.74   | 1.03  | 385 | 7.20   | 0.84 | 361 |
| Egg shell weight at 60 wk (ESW60,g) | 6.96   | 0.66  | 280 | 6.53   | 0.63 | 307 |
| Yolk weight at 40 wk (YW40, g)  | 15.51 | 1.05  | 385 | 14.84 | 1.23 | 361 |
| Yolk weight at 60 wk (YW60, g)  | 17.29 | 1.17  | 275 | 16.99 | 1.68 | 308 |
| Egg shell strength at 40 wk (ESS40, Kg/cm2) | 3.046 | 0.647 | 385 | 2.976 | 0.619 | 361 |
| Egg shell strength at 60 wk (ESS60, Kg/cm2) | 2.888 | 0.612 | 281 | 2.719 | 0.649 | 320 |
| Egg shell thickness at 40 wk (EST40, mm) | 0.316 | 0.024 | 385 | 0.301 | 0.024 | 361 |
| Egg shell thickness at 60 wk (EST60, mm) | 0.329 | 0.028 | 279 | 0.306 | 0.029 | 313 |
| Albumen height at 40 wk (AH40, mm) | 6.2    | 1.0   | 385 | 6.6   | 1.0  | 361 |
| Albumen height at 60 wk (AH60, mm) | 5.8    | 0.9   | 277 | 6.4   | 1.0  | 316 |
| Haugh unit at 40 wk (HU40)      | 79.6  | 8.4   | 385 | 83.2  | 7.7  | 361 |
| Haugh unit at 60 wk (HU60)      | 75.3  | 6.8   | 273 | 79.2  | 7.2  | 316 |

1Phenotypic values of traits in the both two strains are not within the normal distribution.
2Phenotypic values of EST40 and ESW60 in WL are not within the normal distribution.
3Phenotypic values of ESS60 in DW are not within the normal distribution.

doi:10.1371/journal.pone.0028600.t002

As egg weight has a great impact on eggshell and yolk weight, it was taken as a covariant in the association test for ESW and YW. Two significant SNPs were detected by combined analysis for ESW40, one on GGA2 (combined P = 5.85E-09) and the other on GGA3 (combined P = 1.41E-07). The significant SNP on GGA2 is located in the intron2 of odd Oz/ten-m homolog 2 gene (ODZ2). For EN, one SNP at 21.67 Mb on GGA7 with combined P value of 3.97E-07 showed association, and it was located in the intron12 of growth factor receptor-bound protein 14 gene (GRB14).

Egg Shell Weight (ESW) and Yolk Weight (YW)

As egg weight has a great impact on eggshell and yolk weight, it was taken as a covariant in the association test for ESW and YW. Two significant SNPs were detected by combined analysis for ESW40, one on GGA2 (combined P = 5.85E-09) and the other on GGA3 (combined P = 1.41E-07). The significant SNP on GGA2 is located in the intron2 of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 gene (GalNT1). The significant SNP on GGA3 is at site of 100.09 Mb and 21,753 bps away from the nearest known gene BLK. An association signal was also detected in analysis of ESW60, which is at about 9.59 Mb on GGA11 and in the intron3 of zinc finger protein 536 gene (ZNF536). For YW40, only one SNP at 184.63 Mb on GGA1 reached the genome-wise significant level, which is located in ataxia telangiectasia mutated gene (ATM).

Egg Shell Thickness (EST)

Two SNPs located within an 8.12 Mb segment (between 171.22 Mb to 179.35 Mb) on GGA1 were found to be significantly associated with EST40 by using combined P method. One of them (combined P = 9.21E-07) is located in a predicted gene LOC418918, and the other one (combined P = 1.89E-07) is about 22,428 bps away from a known gene ENOX1.

Chromosome-wise Significant SNPs

For other phenotypes including AH and HU, the combined P values of association tests did not reach the genome-wise significant level, but some SNP-trait combinations showed suggestive evidence with chromosome-wise significant level (Table S1), especially with SNPs in QTL regions identified in previous studies, which can provide valuable references for subsequent researches.

Discussion

For domestic animals, the genome-wide association study is becoming a powerful approach for genetic dissection of trait loci along with the completion of genome sequencing and development of high density SNP array. Recently, GWAS has already been applied in the cattle and revealed several loci impacting economically important traits [11,12,13]. Following successful application of GWAS in cattle, we conducted a GWAS in White Leghorn and dwarf layer populations and provided strong evidences for association of SNPs with 7 traits of egg production and quality. A remarkable aspect of our study is that most SNPs exceeded the chromosome-wise significance threshold (Table S2). A few dominance effects reached chromosome-wise significance but were less significant (data not shown). Application of the new chicken SNP array allowed genotyping at a higher marker density than most previous studies, hence some previously undetected loci for egg quality and production traits were found in the present study. The details of all genome-wise significant SNPs, including their positions in the genome, raw P value in each population, combined P values, and candidate genes are summarized in Table 4 and further described as follows.

Age at First Egg (AFE) and Number of Eggs (EN)

A SNP on GGA13 was found to be significantly associated with AFE (combined P = 1.42E-06), and it was located in the intron2 of odd Oz/ten-m homolog 2 gene (ODZ2). For EN, one SNP at 21.67 Mb on GGA7 with combined P value of 3.97E-07 showed association, and it was located in the intron12 of growth factor receptor-bound protein 14 gene (GRB14).

Table 4 and further described as follows.

Egg Weight at 40 wk (EW40, g)1

As egg weight has a great impact on eggshell and yolk weight, it was taken as a covariant in the association test for ESW and YW. Two significant SNPs were detected by combined analysis for ESW40, one on GGA2 (combined P = 5.85E-09) and the other on GGA3 (combined P = 1.41E-07). The significant SNP on GGA2 is located in the intron2 of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 gene (GalNT1). The significant SNP on GGA3 is at site of 100.09 Mb and 21,753 bps away from the nearest known gene BLK. An association signal was also detected in analysis of ESW60, which is at about 9.59 Mb on GGA11 and in the intron3 of zinc finger protein 536 gene (ZNF536). For YW40, only one SNP at 184.63 Mb on GGA1 reached the genome-wise significant level, which is located in ataxia telangiectasia mutated gene (ATM).

Egg Shell Thickness (EST)

Two SNPs located within an 8.12 Mb segment (between 171.22 Mb to 179.35 Mb) on GGA1 were found to be significantly associated with EST40 by using combined P method. One of them (combined P = 9.21E-07) is located in a predicted gene LOC418918, and the other one (combined P = 1.89E-07) is about 22,428 bps away from a known gene ENOX1.

Chromosome-wise Significant SNPs

For other phenotypes including AH and HU, the combined P values of association tests did not reach the genome-wise significant level, but some SNP-trait combinations showed suggestive evidence with chromosome-wise significant level (Table S1), especially with SNPs in QTL regions identified in previous studies, which can provide valuable references for subsequent researches.
Table 3. Phenotypic mean, standard deviation and status of normalization for non-normal traits after the transformation.

| Trait | White Leghorn | Dwarf layer | Status of normalization |
|-------|---------------|-------------|------------------------|
|       | Mean          | SD          | n          | Mean          | SD          | n          |                      |
| EN    | 7.66E05       | 2.42E05     | 385       | 7.37E05       | 2.25E05     | 361       | Yes                  |
| AFE   | 4.38E-05      | 5.12E-06    | 385       | 6.66E-01      | 3.23E-05    | 361       | Yes                  |
| FEW   | 6.29          | 0.39        | 385       | 3.66          | 0.12        | 361       | Yes                  |
| HU60  | 7478.14       | 1465.22     | 273       | 1.71E05       | 4.04E04     | 316       | Yes                  |
| EST40 | 7.18E05       | 1.29E05     | 385       | –             | –           | –         | Yes                  |
| ESW60 | 11.63         | 1.69        | 280       | –             | –           | –         | Yes                  |

1 EST40 and ESW60 have been transformed only for the White Leghorns.
2 ESS60 has been transformed only for the dwarf layers.

doi:10.1371/journal.pone.0028600.t003

Figure 1. Genome-wide scan for egg production and quality traits: –log10 of the combined P value analysis for association with SNPs. Chromosome 1–28 and linkage group E22 are shown in alternating colors for clarity. The horizontal lines indicate the genome-wise significance threshold: –log10 (1.51E-06).
doi:10.1371/journal.pone.0028600.g001
found at genome-wide significance level are within the known genes, indicating that there are disequilibrium between the marker SNP and the causative variation within or near genes, though the characteristics and functions of these genes have not been studied in depth. Identifications of these loci may provide new insights into the genetic basics of egg production and quality traits. Another notable aspect of the present study is that most of the significant SNP were additive with little dominance detected. The reason for this result may be that the dominance variation is less in purebred laying hens than in crossbred ones [23]. Furthermore, for some egg production traits such as EN and EW, dominant variance just presents a small ratio of total phenotypic variance [24]. Similarly little dominance effects with genome-wise significance were detected in our study.

Number of eggs and age at first egg are two important production traits in layers, and producing hens with earlier sexual maturity and higher rate of lay has always been the goal of egg-type chicken breeding. As these reproductive traits are sex-limited and have low to moderate heritability, they would greatly benefit from marker-assisted selection, where the selection can be directed towards actual genetic variation. In this study, a most significant SNP associated with egg number was found to be located in the intron12 of GRB14 gene that encodes a growth factor receptor-binding protein. In human and mammals, GRB14 mRNA was found to be expressed at high level in the ovary, liver, kidney, skeletal muscle and so on [25,26]. It interacts with insulin receptors (IR) and insulin-like growth-factor receptors (IGFR), and may play an inhibiting role in tyrosine kinase receptor (Tkr) signaling pathways [27,28]. IGF and IGFR were reported to regulate ovarian functions and follicular developments in chickens [29,30]. Although the function of GRB14 in chicken is undefined, it may combine with the IGF system to influence egg production in layers.

Age at first egg is an indicator of sexual maturity, and can be impacted by several factors including nutrition, photoperiod and genetics. In the present study, a SNP in the intron2 of OD22 gene was revealed to be associated significantly with Age at first egg. OD22, also known as Teneurin-2, encodes a neuronal cell surface protein and plays an important role in development of nervous system [31]. It was found that Teneurin are expressed prominently in developing chicken brain, and especially in the visual system including retina and optic tectum [32]. The current study provides the first report that Teneurin-2 may have effect on the sexual maturity of chickens. A recent study found that expressions of genes in the nervous system can influence the age when chickens lay their first egg [33]. Furthermore, some previous studies revealed that light intensity can influence layers’ age at first egg and longer light periods can lead to earlier sexual maturity [34,35]. As stimulations of lights play roles mainly through the visual and nervous system, genes related to these systems may impact sexual maturity in chickens.

In addition to egg production, egg quality is another major selection criterion in poultry breeding, especially eggshell quality. Good eggshell quality is not only important for reproductive performance but also for human consumption. Eggshell weight, eggshell strength, and eggshell thickness are important indications of eggshell quality. We identified several significant SNPs influencing eggshell weight at different age. One significant SNP associated with ESW40 is in the intron2 of GALNT1 gene. In human, some nucleotide mutations of GALNT1 may cause ovarian cancer [36]. On the other hand, normal GALNT1 may ensure normal functions of ovary. The characteristic of this gene is still unclear in chickens, and the current study is the first report that its polymorphism is associated with egg quality traits. Another ESW40 associated SNP with genome-wise significance is on GGA2, where there are also three chromosome-wise significant SNPs. These four SNPs are located in a 645 Kb segment that may be a novel QTL as it does not coincide with previously reported QTL or candidate gene for ESW [37,38,39]. In this putative QTL region, there are lots of known genes including genes related to DNA modification, transcription, replication and RNA translation (NEIL2, GATA4, MCM5 and TRAM2); genes related to immune system (IL17, Antimicrobial peptide CHP1 and beta-defensin gene cluster); a gene plays a role in the calcium homeostasis (EFH1). Functions of most genes mentioned above are not fully understood in chickens and need to be studied further, although they have been studied extensively in human.

Significant SNPs differed over time for egg shell weight, and a SNP on GGA11 was found to be significantly associated with ESW60. Fairfull and Gowe noted high correlations over ages in egg quality traits [40], but Abashit et al. found that in selected lines, similar populations as used in the present study, correlations between early and late traits were weak [41]. Therefore, it is suggested that the partial genetic independence exists in some egg quality traits at different ages and identifying QTL that differ over time for the same trait is important. The significant SNP associated with ESW60 also showed suggestive association (chromosome-wise significance) with EST60 (combined $P=2.04E-06$) and ESS40 (combined $P=1.72E-06$), and it is located in the intron3 of $\xi NF536$ gene that encodes a kind of DNA binding protein and functions as a transcriptional repressor [42,43]. This is the first report that $\xi NF536$ may affect egg shell weight in chickens.

**Table 4. Genome-wise significant ($P<1.51E-06$, Bonferroni correction) SNPs for egg production and quality traits.**

| SNP          | Associated Traits GGA | Position (bp) | $^1$Pwl | $^1$Pdw | $^2$Combined P | Candidate gene |
|--------------|------------------------|---------------|---------|---------|---------------|----------------|
| GaluGA315030 | EN                     | 7             | 21676854| 6.17E-01| 3.45E-08      | 3.97E-07       | GRB14         |
| GaluGA092322 | AFE                    | 13            | 4796267 | 6.85E-06| 1.19E-02      | 1.42E-06       | OD22          |
| rs13636444   | ESW40                  | 2             | 86114050| 6.92E-10| 3.66E-01      | 5.85E-09       | GALNT1        |
| rs14411524   | ESW40                  | 3              | 110095288| 1.40E-07| 5.09E-02      | 1.41E-07       | BLK           |
| rs14022717   | ESW60                  | 11            | 9596922 | 2.24E-07| 2.49E-01      | 8.62E-07       | ZNF536        |
| GaluGA059301 | YW40                   | 1             | 184632695| 3.31E-02| 1.1E-06       | 2.56E-08       | ATM            |
| rs13968878   | EST40                  | 1             | 171224927| 6.91E-08| 1.41E-01      | 2.81E-08       | ENOX1          |
| rs13978498   | EST40                  | 1             | 179350984| 1.74E-02| 2.97E-06      | 9.22E-07       | LOC418918     |

$^1$Results of independent association test in White Leghorn and dwarf layer populations.

$^2$Combined P value follow the Fisher’s method.

doi:10.1371/journal.pone.0028600.0004
Lots of QTL regions affecting eggshell thickness have been detected by previous linkage studies and they distribute on GGA1, GGA2, GGA3, and GGA7 [37,38,44]. Some candidate genes for eggshell thickness were also identified on GGA2, GGA4, GGA8 and GGA9 [5,6]. In this study, two association signals were found on GGA1 for EST40, which is located in a hypothetical locus LOC418918 and the other in a known gene EVOX1. The region harboring these two SNPs ranges from 171.22 Mb to 179.53 Mb, which may be a novel QTL for eggshell thickness and about 70 Mb away from the QTL reported by Sasaki et al. [38].

Internal quality is becoming another focus of attention in improving egg quality traits, especially as egg consumption is changing from shell eggs toward egg products. Yolk weight is an important indication of internal egg quality. In this study, an association signal at genome-wise significant level was found for YW40, located in ATP gene on GGA1. ATP can regulate a wide variety of downstream proteins mainly related to genome stabilizing and cell cycle controlling [43,46,47]. This is an important candidate gene, but in chickens its function is still unclear and needs future study.

In order to avoid the extreme conservation induced by Bonferroni correction and increase probabilities of finding potential genetic variant impacting egg production and quality traits, association tests were also performed for markers on each chromosome individually. Some chromosome-wise significant SNPs were found in locations within the previously reported QTL regions. On GGA4, there were four QTL regions affecting egg weight, and their range span 51.6–52.6 Mb, 46.7–46.8 Mb, 70.9–80.3 Mb and 61.5–81.3 Mb [38,48,49]. In the present study, three SNPs associated with EW40 chromosome-wise are within a segment ranging from 78.67 Mb to 79.36 Mb, which coincides with the QTL regions reported by Sasaki et al. and Tuiskula-Haavisto et al. [38,48]. For FEW, the chromosome-wise significant SNPs are located in a QTL region (38.00 Mb to 38.53 Mb) on GGA1 for egg weight at 29 wk in a red junglefowl (RJF) x White Leghorn (WL) cross reported by Kerje et al. [49].

As large differences of genetic background and phenotypic distribution exist between White Leghorn and dwarf layers, data from individual studies of these two populations cannot be analyzed together in a single association test. Otherwise it would induce false positive caused by population stratification. Therefore, we conducted association tests separately and then performed a meta-analysis to increase the statistical power in estimating the true effect signals. Fisher’s combined probability method, although proposed decades ago, is a simple but elegant technique for meta-analysis. This method is appropriate to combine the results from several independent tests bearing upon the same overall hypothesis. For the present study, association tests in two different experimental populations were under the same null hypothesis ($H_0$): no SNP associated with the trait. Furthermore, as the basis of the combined probability method should be one-sided test, and in order to avoid the heterogeneity, we rejected SNP markers with opposing direction of effect in separate studies, and combined ones with additive effect in the same direction. Therefore, Fisher’s method is particularly well suited as a meta-analysis tool for this study that performed association tests across different experimental populations.

The number of identified SNPs is limited and these loci might not fully describe genetic diversity underlying traits in this study. Furthermore, the genetic mechanisms of quantitative traits might involve complex interactions among genes and between genes and environmental conditions, or epigenetic mechanisms which cannot be captured by additive models. Therefore, increasing density of markers in the genotyping panels and improving genetic models and statistical methods may benefit the detection of causative genetic variability for quantitative traits in domestic animals.

In summary, the current study revealed 8 genome-wise significant and 95 chromosome-wise significant SNPs for egg production and quality traits in White Leghorn and brown-egg dwarf layers by using the high-density SNP array and association analysis based on least squares regression. Some SNPs are located in possible causative genes or within the previously reported QTL region, but most of the significant SNPs are reported, for the first time, to be associated with egg production and quality traits. To our knowledge, this is the first publication of GWAS on egg production and quality traits in chickens and our findings lay a preliminary foundation for follow-up studies to identify causal mutations by enriching markers within the identified intervals and functional studies on related genes, and subsequently they may be applied in marker-assisted selection program on egg layers.

Supporting Information

Table S1 Chromosome-wise significant threshold for each chromosome. (DOC)
Table S2 Chromosome-wise significant trait-SNP combinations. (DOC)

Acknowledgments

The authors are grateful to Dr. Yang Da at the University of Minnesota for providing us with the statistical software EPISNP and helpful discussion on the current research.

Author Contributions

Conceived and designed the experiments: NY. Performed the experiments: DL, LI, JL. Analyzed the data: DL, LI, JL. Contributed reagents/materials/analysis tools: SC, LQ, JZ, GX. Wrote the paper: WL, JL, NY.

References

1. Dekkers JC (2004) Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. J Anim Sci 82 E-Suppl: ES13-320.
2. Lamont SJ, Lakshmanan N, Photos Y, Kaiser MG, Kuhn M, et al. (1996) Genetic markers linked to quantitative trait loci in poultry. Anim Genet 27: 1–8.
3. Abash B, Dekkers JC, Lamont SJ (2006) Review of quantitative trait loci identified in the chicken. Poult Sci 85: 2079–2096.
4. Bhattacharya TK, Chatterjee RN, Sharma RP, Niranjana M, Rajkumar U (2010) Associations between novel polymorphisms at the 5'UTR region of the prolactin gene and egg production and quality in chickens. Theriogenology.
5. Dunn IC, Joseph NT, Bain M, Edmond A, Wilson PW, et al. (2009) Polymorphisms in eggshell organic matrix genes are associated with eggshell quality measurements in Red Junglefowl. Anim Genet 41: 110–114.
6. Yao JF, Chen ZX, Xu GY, Wang XL, Ning ZH, et al. (2010) Low-density lipoprotein receptor-related protein 8 gene association with egg traits in dwarf chickens. Poult Sci 89: 853–856.
7. Wong GK, Liu B, Wang J, Zhang Y, Yang X, et al. (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432: 717–722.
8. Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, et al. (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science 328: 532–538.
9. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, et al. (2009) Development and characterization of a high density SNP genotyping assay for cow. PLoS One 4: e3350.
10. Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.
11. Jiang L, Liu J, Sun D, Ma P, Ding X, et al. (2010) Genome-wide association studies for milk production traits in Chinese Holstein population. PLoS One 5: e13661.
12. Sahana G, Guldbrandtsen B, Bendixen C, Lund MS (2010) Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. Anim Genet 41: 579–588.
13. Bolormaa S, Hayes BJ, Savin K, Hawkins R, Barendse W, et al. (2011) Genome-wide association studies for feedlot and growth traits in cattle. J Anim Sci.

14. Zhang LC, Ning ZH, Xu GY, Hou ZC, Yang N (2005) Heritabilities and genetic and phenotypic correlations of egg traits in brown-egg dwarf layers. Poult Sci 84: 1209–1213.

15. Yang N, Ning Z, Wang Q, Li J, Wu C (1996) Application of the sex-linked dwarf gene to improving feed efficiency of layers. In: Proceedings of 20th World's Poultry Congress, New Delhi, India. pp 447–451.

16. Ma L, Runeda HB, Dvorkin D, Garbe JR, Da Y (2008) Parallel and serial computing tools for testing single-locus and epistatic SNP effects of quantitative traits in genome-wide association studies. BMC Bioinformatics 9: 315.

17. Mao Y, Loudon NR, Ma L, Dvorkin D, Da Y (2006) Detection of SNP epistatic effects of quantitative traits using an extended Kempthorne model. Physiol Genomics 28: 46–52.

18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.

19. Arnott DK, Meyers KJ, Devereux RB, Tiwari HC, Gus CG, et al. (2011) Genetic variation in NCAM1 contributes to left ventricular wall thickness in hypertensive families. Circ Res 108: 279–283.

20. Carraquillo MM, Beltin O, Hunter TA, Ma L, Bisciglio GD, et al. (2011) Replication of BDNF Association with Alzheimer's Disease and Evaluation of Genetic Interactions. J Alzheimers Dis.

21. Hess A, Iyer H (2007) Fisher's combined p-value for detecting differentially expressed genes using Affymetrix expression arrays. BMC Genomics 8: 96.

22. Han B, Kang HM, Eskin E. (2009) Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS Genet 5: e1000456.

23. Breshe B, Gibson JP (1999) Genetic variation of egg production traits in purebred and crossbred laying hens. Anim Sci 68: 433–439.

24. Wei M, van der Weer JH (1993) Animal model estimation of additive and dominance variances in egg production traits of poultry. J Anim Sci 71: 57–65.

25. Daly RJ, Sanderson GM, James PW, Sutherland RL (1996) Cloning and characterization of GRB14, a novel member of the GRB7 gene family. J Biol Chem 271: 12502–12510.

26. Smith TP, Grosse WM, Freking BA, Roberts AJ, Stone RT, et al. (2001) Inhibition of insulin receptor catalytic activity by the molecular adapter Grb14. J Biol Chem 277: 4845–4852.

27. Hemmings R, Agatep R, Badani K, Wiant K, Arthur G, et al. (2001) Human growth factor receptor bound 14 binds the activated insulin receptor and alters the insulin-stimulated tyrosine phosphorylation levels of multiple proteins. Biochem Cell Biol 79: 21–32.

28. McMurtry JP, Francis GL, Upton Z (1997) Insulin-like growth factors in poultry. In: Poultry Breeding and Genetics Crawford RD, ed. Amsterdam, Netherlands: Elsevier. pp 705-759.

29. Ashish B, Sandford E, Arango J, Settar P, Fulton JE, et al. (2009) Extent and consistency of linkage disequilibrium and identification of DNA markers for production and egg quality traits in commercial layer chicken populations. BMC Genomics 10 Suppl 2: S2.

30. Strausberg RL, Feingold EA, Grous HE, Derge JG, Klausner RD, et al. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 99: 16899–16903.

31. Nakai T, Hino K, Wada S, Maeda H (2003) Identification of the DNA binding specificity of the human ZNF219 protein and its function as a transcriptional repressor. DNA Res 10: 135–145.

32. Wright D, Keje S, Lundstrom K, Baehl J, Schutz K, et al. (2006) Quantitative trait loci analysis of egg and meat production traits in a red junglefowl x White Leghorn cross. Anim Genet 37: 529–534.

33. Kastan MB, Lane DS (2000) The many substrates and functions of ATM. Nat Rev Mol Cell Biol 1: 179–186.

34. Kurz EU, Lees-Miller SP (2004) DNA damage-induced activation of ATM and ATM-dependent signaling pathways. DNA Repair (Amst) 3: 889–900.

35. Negrini S, Gorgoulis VG, Halazonetis TD (2010) Genomic instability—an evolving hallmark of cancer. Nat Rev Mol Cell Biol 11: 220–229.

36. Tsuchiya-Haavisto M, Honkatiemi M, Villki J, de Koning DJ, Schulman NF, et al. (2002) Mapping of quantitative trait loci affecting quality and production traits in egg layers. Poult Sci 81: 919–927.

37. Keje S, Carlborg O, Jacobsson L, Schutz K, Hartmann C, et al. (2003) The twofold difference in adult size between the red junglefowl and White Leghorn chickens is largely explained by a limited number of QTLs. Ani Genet 34: 264–274.