

RESEARCH ARTICLE

Open Access



# Genome-wide association study revealed a promising region and candidate genes for eggshell quality in an F<sub>2</sub> resource population

Congjiao Sun<sup>1†</sup>, Liang Qu<sup>2†</sup>, Guoqiang Yi<sup>1</sup>, Jingwei Yuan<sup>1</sup>, Zhongyi Duan<sup>1</sup>, Manman Shen<sup>2</sup>, Lujiang Qu<sup>1</sup>, Guiyun Xu<sup>1</sup>, Kehua Wang<sup>2</sup> and Ning Yang<sup>1\*</sup>

## Abstract

**Background:** Eggshell is subject to quality loss with aging process of laying hens, and damaged eggshells result in economic losses of eggs. However, the genetic architecture underlying the dynamic eggshell quality remains elusive. Here, we measured eggshell quality traits, including eggshell weight (ESW), eggshell thickness (EST) and eggshell strength (ESS) at 11 time points from onset of laying to 72 weeks of age and conducted comprehensive genome-wide association studies (GWAS) in 1534 F<sub>2</sub> hens derived from reciprocal crosses between White Leghorn (WL) and Dongxiang chickens (DX).

**Results:** ESWs at all ages exhibited moderate SNP-based heritability estimates (0.30 ~ 0.46), while the estimates for EST (0.21 ~ 0.31) and ESS (0.20 ~ 0.27) were relatively low. Eleven independent univariate genome-wide screens for each trait totally identified 1059, 1026 and 1356 significant associations with ESW, EST and ESS, respectively. Most significant loci were in a region spanning from 57.3 to 71.4 Mb of chromosome 1 (GGA1), which together account for 8.4 ~ 16.5 % of the phenotypic variance for ESW from 32 to 72 weeks of age, 4.1 ~ 6.9 % and 2.95 ~ 16.1 % for EST and ESS from 40 to 72 weeks of age. According to linkage disequilibrium (LD) and conditional analysis, the significant SNPs in this region were in extremely strong linkage disequilibrium status. Ultimately, two missense SNPs in GGA1 and one in GGA4 were considered as promising loci on three independent genes including *ITPR2*, *PIK3C2G*, and *NCAPG*. The homozygotes of advantageously effective alleles on *PIK3C2G* and *ITPR2* possessed the best eggshell quality and could partly counteract the negative effect of aging process. *NCAPG* had certain effect on eggshell quality for young hens.

**Conclusions:** Identification of the promising region as well as potential candidate genes will greatly advance our understanding of the genetic basis underlying dynamic eggshell quality and has the practical significance in breeding program for the improvement of eggshell quality, especially at the later part of laying cycle.

**Keywords:** GWAS, Eggshell quality, Age-dependent traits, SNP-based heritability, Chicken

\* Correspondence: nyang@cau.edu.cn

†Equal contributors

<sup>1</sup>National Engineering Laboratory for Animal Breeding and MOA Key Laboratory of Animal Genetics and Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China  
Full list of author information is available at the end of the article

## Background

A vast number of eggs are produced annually for human consumption worldwide. Avian calcified eggshell as a unique bioceramic material can provide protection to egg contents from physical damage and microorganic invasion [1, 2]. Changes in eggshell properties are directly related to increasing risk of foodborne disease for the consumer [3]. Low eggshell quality will also lead to more cracked eggs in the automatic sorting and packing process in the modern egg industrial production [4]. It has been estimated that improving the mean eggshell strength by one Newton will lead to 0.5 % less broken eggs per hen per laying cycle [5]. On the other hand, eggshell is biologically significant for bird embryo developments by allowing gas exchange between the embryo and the environment and supplying calcium for the embryo bone deposition [6, 7]. Hatching eggs with thin eggshell have high embryonic mortality, owing to more loss of water vapour during the incubation [8]. Furthermore, eggshell are subject to quality loss with the aging process of laying hens [9], which hinders the developmental trend to prolong the laying cycles of egg-type chicken in the future. Therefore, understanding the genetic control for dynamic eggshell quality with aging process is of great economic and biological importance.

In recent years, the genomic [5, 10], transcriptomic [11–13], proteomic [14–18] and structural analyses [19–21] of eggshell have been carried out to provide new insights into better understanding on the eggshell mineralization and its ultrastructure or microstructure that contribute to eggshell quality. Identifying the quantitative trait loci (QTLs) that relate to eggshell quality is one of the most powerful strategies to illustrate the genetic mechanism underlying eggshell quality. Up to date, a total of 62 QTLs related to eggshell quality have been collected in the AnimalQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/index>). However, most previously reported QTLs were identified by low-density linkage analysis with limited markers of microsatellites [22–24], which restricted the confidence intervals and mapping accuracy [25, 26]. Recently, genome-wide association study (GWAS) has been utilized to identify the associations between genomic loci and phenotypes with relatively high-density SNP arrays in chicken [27]. For instance, Liu et al. [10] conducted the first GWA analysis with relatively high-density SNA array (60 K) to refine the associations with egg production and egg quality traits in chicken. With the rapid advance in next-generation sequencing technologies, large amounts of SNPs in chicken have been discovered [28]. The development of 600 K Affymetrix Chicken SNP array allows more efficient screening of the causal loci and genes in relevance to target traits.

It is notable that many complex traits are dynamic varied with the aging process of animals [29, 30]. However, previous studies utilized the phenotypes from limited age points. Animal phenotypes at different age should be utilized in GWA analysis to refine the associations with age-dependent complex traits and increase the statistical power. Growing evidences had been provided that this type of “longitudinal design” could efficiently identify the time-dependent or consistent loci for complex traits [31, 32].

In the current study, we conducted GWA analysis on the dynamic eggshell quality traits at 11 time points from onset of laying to 72 weeks old utilizing a 600 K high-density SNP arrays in an  $F_2$  chicken population. We aimed to explore the associated genomic loci and genes that contribute to the phenotypic variability and dynamic change in eggshell quality traits, and to provide potential breeding tools for the improvement of eggshell quality.

## Results

### Phenotype description and genetic parameters

Means and standard deviations for eggshell quality traits, including ESW, EST and ESS, at 11 time points from onset of laying to 72 weeks of age are presented in Table 1 and Additional file 1: Figure S1. These three eggshell quality traits displayed a smooth curve with hen age except 36 weeks-old, at which ESW, EST and ESS showed a abrupt decrease (Additional file 1: Figure S1) and the results of GWA analysis in this week were also unusual (Table 3). We speculated that this might be caused by the feed or environment changes in this period leading to eggshell quality decrease. For this reason, phenotypes collected at

**Table 1** Descriptive statistics of eggshell traits in the  $F_2$  population

Age (wks)	N	Eggshell traits (Mean $\pm$ SD)		
		ESW (g)	EST (mm)	ESS (kg/cm <sup>2</sup> )
AFE	1494	3.69 $\pm$ 0.36	0.393 $\pm$ 0.017	3.61 $\pm$ 0.57
32	1473	4.40 $\pm$ 0.45	0.402 $\pm$ 0.021	3.70 $\pm$ 0.67
36	1464	4.37 $\pm$ 0.55	0.396 $\pm$ 0.022	3.59 $\pm$ 0.75
40	1476	4.64 $\pm$ 0.50	0.401 $\pm$ 0.021	3.67 $\pm$ 0.72
44	1420	4.63 $\pm$ 0.55	0.399 $\pm$ 0.023	3.60 $\pm$ 0.81
48	1226	4.70 $\pm$ 0.59	0.401 $\pm$ 0.024	3.51 $\pm$ 0.86
52	1225	4.68 $\pm$ 0.59	0.396 $\pm$ 0.024	3.43 $\pm$ 0.86
56	1348	4.74 $\pm$ 0.61	0.395 $\pm$ 0.024	3.40 $\pm$ 0.89
60	1364	4.70 $\pm$ 0.65	0.392 $\pm$ 0.024	3.25 $\pm$ 0.92
66	1304	4.62 $\pm$ 0.64	0.387 $\pm$ 0.028	3.17 $\pm$ 0.93
72	1253	4.57 $\pm$ 0.62	0.395 $\pm$ 0.025	3.02 $\pm$ 0.80

N number of samples, ESW eggshell weight, EST eggshell thickness, ESS eggshell strength, Mean arithmetic mean, SD standard deviation, AFE age of first egg

**Table 2** Summary of genetic analysis for eggshell weights at different wks of age

Traits <sup>a</sup>	FESW	ESW32	ESW36	ESW40	ESW44	ESW48	ESW52	ESW56	ESW60	ESW66	ESW72
FESW	0.41(0.04)	0.87(0.04)	0.87(0.05)	0.77(0.06)	0.83(0.06)	0.77(0.06)	0.77(0.06)	0.66(0.07)	0.70(0.07)	0.68(0.07)	0.65(0.07)
ESW32	0.53	0.30(0.04)	1.00(0.03)	0.99(0.02)	0.96(0.03)	0.93(0.03)	0.96(0.03)	0.87(0.05)	0.90(0.04)	0.89(0.04)	0.87(0.05)
ESW36	0.41	0.57	0.41(0.04)	1.00(0.03)	0.98(0.04)	0.98(0.04)	0.99(0.04)	0.92(0.05)	1.00(0.04)	0.92(0.05)	0.93(0.05)
ESW40	0.46	0.61	0.55	0.37(0.04)	1.00(0.02)	0.99(0.02)	0.97(0.03)	0.92(0.04)	0.95(0.03)	0.96(0.03)	0.92(0.03)
ESW44	0.39	0.61	0.51	0.64	0.46(0.04)	1.00(0.02)	1.00(0.02)	0.98(0.03)	0.98(0.03)	0.96(0.03)	0.96(0.03)
ESW48	0.38	0.57	0.51	0.61	0.64	0.45(0.05)	1.00(0.02)	0.98(0.03)	1.00(0.03)	0.98(0.03)	0.95(0.03)
ESW52	0.40	0.58	0.52	0.62	0.65	0.69	0.39(0.05)	1.00(0.02)	0.99(0.03)	0.97(0.03)	0.95(0.03)
ESW56	0.35	0.52	0.45	0.58	0.61	0.62	0.65	0.33(0.05)	1.00(0.03)	0.95(0.03)	0.93(0.03)
ESW60	0.33	0.49	0.44	0.56	0.54	0.57	0.63	0.62	0.44(0.04)	0.98(0.03)	0.96(0.03)
ESW66	0.36	0.54	0.47	0.58	0.59	0.62	0.65	0.66	0.64	0.40(0.05)	0.99(0.02)
ESW72	0.38	0.56	0.47	0.61	0.63	0.61	0.68	0.67	0.63	0.74	0.43(0.05)

Diagonal: heritability estimates. Lower triangle: phenotypic correlations. Upper triangle: genetic correlations. Standard errors of the estimates are in parentheses  
<sup>a</sup>FESW eggshell weight of first egg, ESW32, ESW36, ESW40, ESW44, ESW48, ESW52, ESW56, ESW60, ESW66, ESW72 eggshell weights of 32,36,40,44,48,52,56,60,66,72 weeks of age

36 weeks-old were not included in the subsequent multi-variate GWA analysis. The additive genetic variations of ESW, EST and ESS at different age were estimated from all eligible markers using GCTA. ESW at all ages exhibited moderate SNP-based heritability estimates (0.30 ~ 0.46, Table 2), and the highest SNP-based heritability estimate was found in ESW44 ( $h_{SNP}^2 = 0.46$ ). However, the estimates for EST (0.21 ~ 0.31, Additional file 2: Table S1) and ESS (0.20 ~ 0.27, Additional file 2: Table S2) were relatively

**Table 3** Number and distribution of significant SNPs for eggshell traits

Age	ESW		EST		ESS
	GGA1	GGA4	GGA1	GGA1	GGA1
AFE	0	1	0	0	0
32	590	2	0	0	230
36	48	7	0	0	121
40	867	3	501	0	809
44	814	0	579	0	624
48	569	0	167	0	378
52	585	0	637	0	596
56	601	0	618	0	608
60	615	0	722	0	870
66	717	0	714	0	1016
72	633	0	721	0	1091
total (Region)	1049 (60.1 ~ 69.0 M)	8	1026 (59.7 ~ 70.4 M)	1356 (57.3 ~ 71.4 M)	
Multi-analysis <sup>a</sup> (Region)	503 (62.4 ~ 67.9 M)	0	501 (60.8 ~ 67.8 M)	532 (60.5 ~ 67.0 M)	

AFE age of first egg, ESW eggshell weight, EST eggshell thickness, ESS eggshell thickness

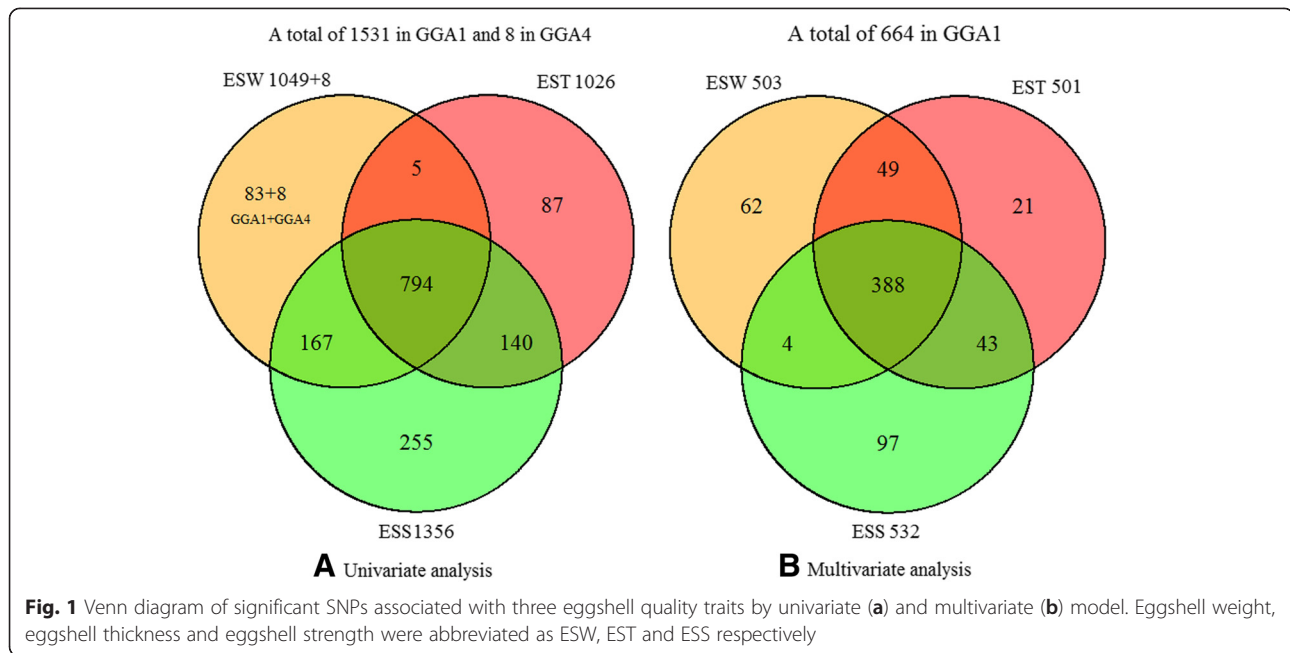
<sup>a</sup>Multi-analysis means using multivariate model for GWA analysis. For the analysis of ESW and ESS, time points of 32 weeks, and 40–72 weeks were used; for EST, time points of 40–72 weeks were used

low. Moreover, bivariate GCTA analyses indicated that ESW at different age were highly and positively correlated, specially at neighboring time points ( $r_g > 0.90$ ).

#### Identification of candidate loci by GWAS

Eggshell quality traits were collected at 11 time points from onset of laying to 72 weeks. At each time point, three separate genome-wide association tests were conducted for each eggshell quality trait with univariate method. A total of 1057, 1026 and 1356 genome-wide significant associations were identified with ESW, EST and ESS, respectively (Table 2, Additional file 2: Table S3). Almost all the significant locus were in a region spanning from 57.3 to 71.4 M of chromosome 1 (GGA1) (Table 3). Out of these loci in this region, 794 SNPs were observed in association with all these three eggshell quality traits at genome-wide significance level (Fig. 1). It was notable that most of the SNPs in this region possess a MAF more than 0.3. In addition, 8 loci on chromosome 4 (GGA4) were significantly associated with ESW, but only for young hens from age of first egg (AFE) to 40 week old. The global view of the putative  $P$ -values for all SNPs affecting eggshell quality traits at 44weeks is given in Fig. 2, and the Manhattan and quantile-quantile (QQ) plots for the remaining time points in Additional file 1: Figure S2.

To enhance the statistical power, a joint GWA analysis of multi time points was performed by fitting these SNPs into a multivariate model. The time points that had significant SNPs were included in this model except for 36 weeks owing to the abnormal phenotype in this week. The 1531 significant SNPs related to eggshell quality identified by univariate method in GGA1 were further analyzed in this multivariate model. Consequently, a total of 503, 501 and 532 significant hits on GGA1 from 60.5 to 67.9 M presented consistent and compelling associations with



longitudinal eggshell quality traits (Table 3, Additional file 2: Table S4), and 388 SNPs out of these locus had pervasive effect on ESW, EST and ESS (Fig. 1b).

#### Linkage disequilibrium (LD) and conditional analysis

The linkage disequilibrium (LD) analysis were then carried out and the results showed the uncovered SNPs in GGA1 were in extremely strong linkage disequilibrium status, especially for the LD block from 64.0 to 67.5 Mb that include half (732/1531) of the total loci (Fig. 3). To detect the independent acting locus in this region, we then carried out stepwise conditional GWA analysis to prioritize separately associated SNPs. At locus *rs312347405*, a missense mutation in association with all three eggshell traits was then fitted into the multivariate model using its genotype as covariate to explore the change of the associations with other loci. As shown in Fig. 3, if the genotype of *rs312347405* was used as covariate in multivariate analysis, the significance levels of all other loci were substantially attenuated below genome-wide significant level. Regional plots and conditional analysis in multivariate model for the traits of eggshell thickness (EST) and eggshell strength (ESS) were shown in Additional file 1: Figure S3.

#### Estimation of contribution to phenotypic variation (CPV)

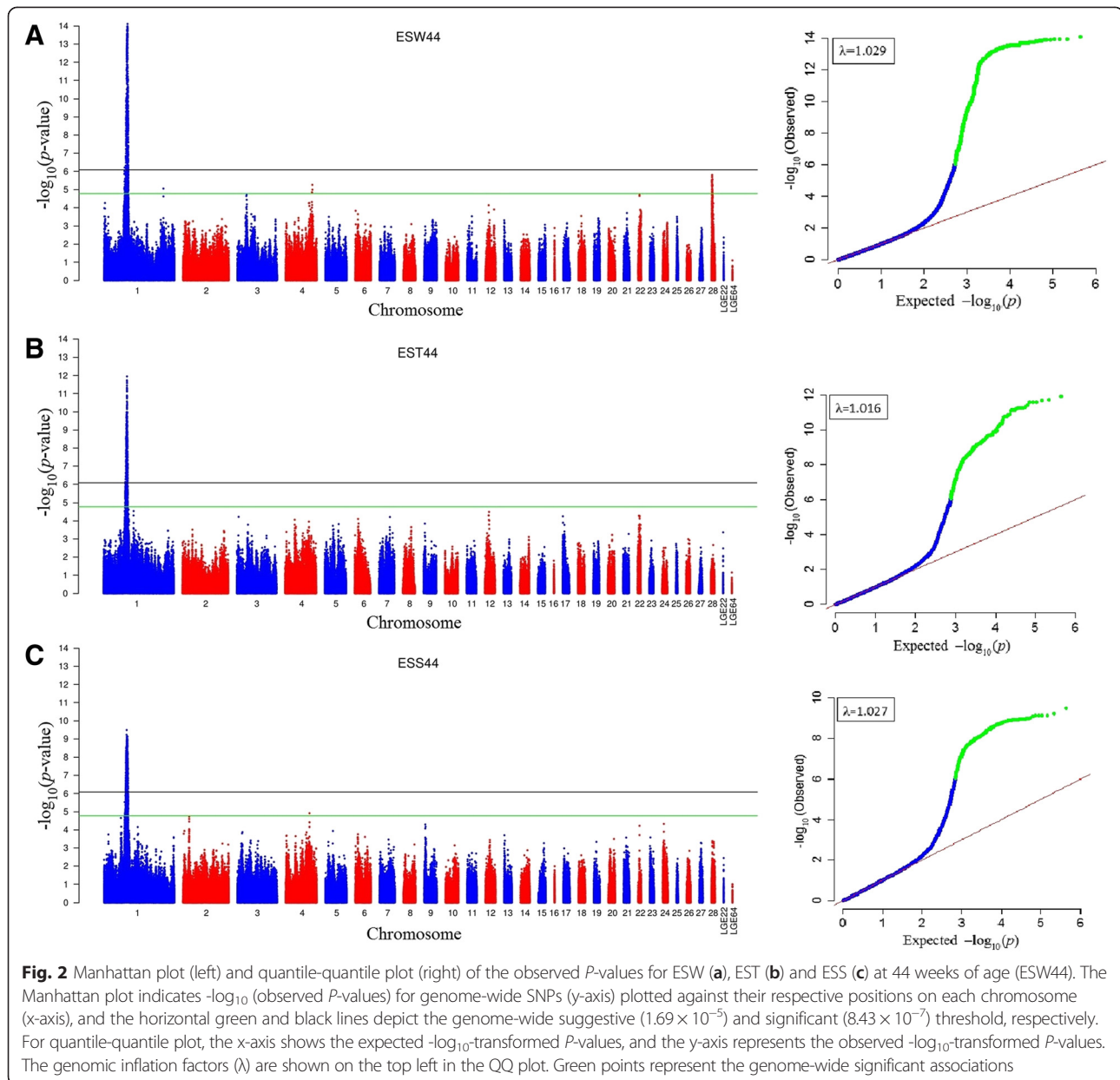
The phenotypic variance explained by loci or genomic region were estimated by a tool of GCTA for three eggshell traits. All together, 1531 significant loci from 57.3 to 71.4 Mb in GGA1 could account for 8.4 ~ 16.5 % of the phenotypic variance for ESW from 32 to 72 weeks of

age, 4.1 ~ 6.9 % and 2.95 ~ 16.1 % for EST and ESS from 40 to 72 weeks of age. For single locus, five missense mutations out of these 1531 loci were estimated for their CPVs (Table 4). Among these five loci, two SNPs presented high CPVs for eggshell traits. One loci, *rs312347405*, could independently explain 5.25–8.7 %, 4.22–7.41 % and 3.18–7.59 % of the phenotypic variance for ESW, EST and ESS from 32 to 72 weeks of age respectively. Notably, the effect alleles at *rs314058619* were associated with ESW at each time point except the first egg. The other one, *rs316607577*, could account for 4.78–7.77 %, 3.42–5.31 % and 2.43–11.52 % of the phenotypic variance respectively.

The phenotypic differences among 3 genotypes at these two loci are displayed in Fig. 4. There was no difference in eggshell quality traits among different genotypes for the first egg. However, along with the aging process, profound differences in eggshell quality appeared among three genotypes. Eggshell of each genotype exerted different rate of quality decay with aging process, i.e., the homozygotes of advantageously effective alleles possessed the best eggshell quality and could partly counteract the negative effect of aging process. However, the change trends of eggshell traits among three genotypes were concordant along with age.

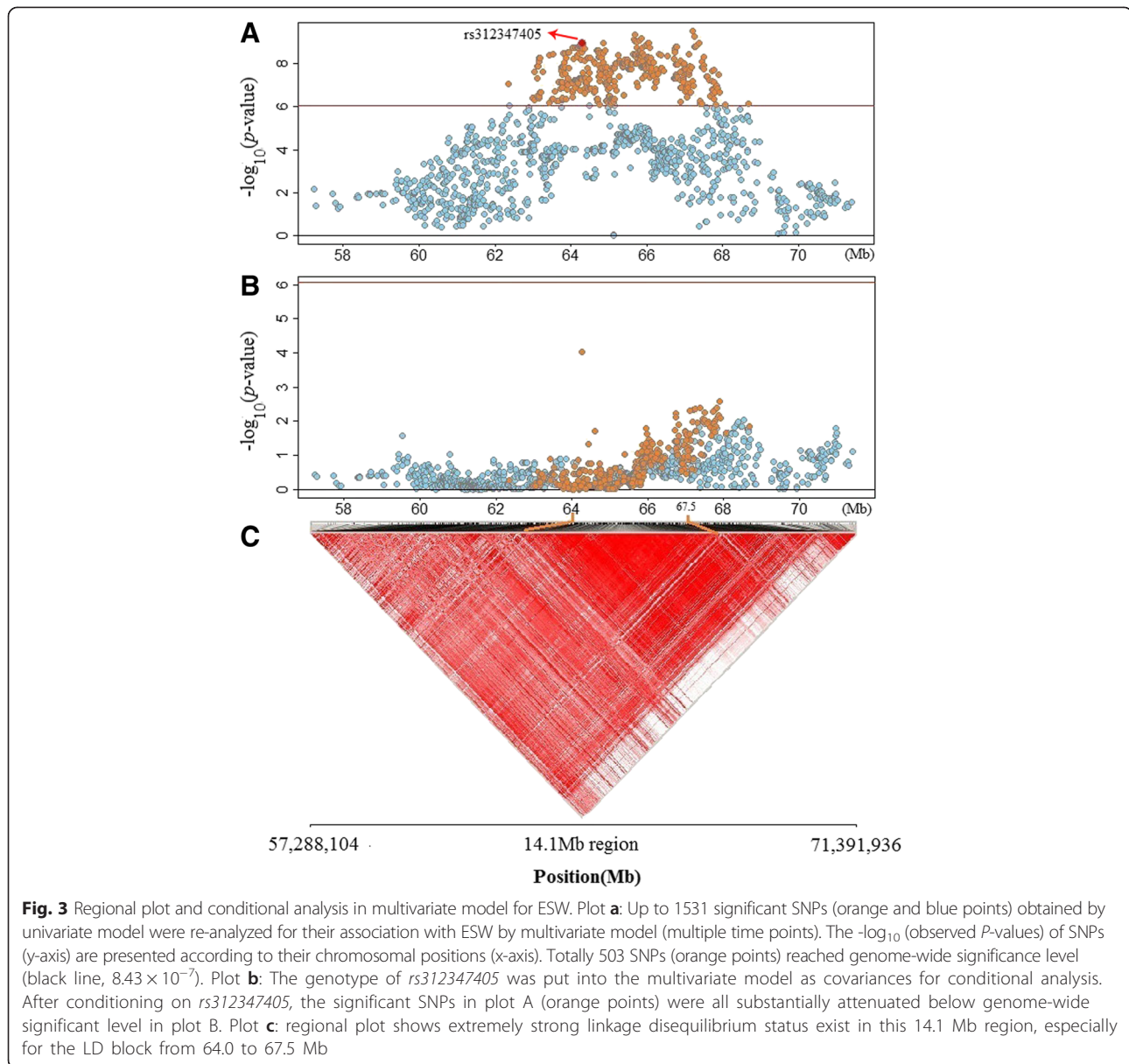
#### Promising genes related to eggshell quality

Utilizing gene annotation of the causal locus allowed us to screen the putative genes relating to eggshell quality. For SNPs-traits association analysis, the missense mutations on exons were more meaningful than



the SNPs located on introns or intergenic regions. We totally found five missense loci at GGA1 and one at GGA4 using univariate genome-wide screens. They located on 6 independent genes including phosphatidylinositol-4-phosphate 3-kinase, catalytic subunit type 2 gamma (*PIK3C2G*), inositol 1,4,5-trisphosphate receptor type 2 (*ITPR2*), RecQ helicase-like (*RECQL*), ATP-binding cassette subfamily C member 9 (*ABCC9*) and cancer susceptibility candidate 1 (*CASC1*) on GGA1 and non-SMC condensin I complex subunit G (*NCAPG*) on GGA4. Nonetheless, only two SNPs, *rs312347405* and *rs316607577*, located on *PIK3C2G* and *ITPR2*, remained significantly associated with eggshell quality traits after multivariate GWA analysis.

Furthermore, as illustrated above, these two SNPs could explain more phenotypic variance than others. Consequently, we first considered *PIK3C2G* and *ITPR2* as the primary candidate genes associated with eggshell quality. The missense mutation in GGA4 that located on *NCAPG* was posterior putative locus for its roles in affecting early eggshell weight. In addition to missense mutations on exons, the mutations on 3' or 5' UTRs were also of biological significance for genetic variation underlying complex traits. We aggregately discovered 15 mutations that located on 3' or 5' UTRs corresponding to 13 independent genes (Additional file 2: Table S5). According to functional annotation and database research, however, most of these



genes had no direct relevance to eggshell calcification or calcium metabolism, except for *rs316447591* which located on the 3'UTR of *IITPR2* gene. The detailed message of promising loci and the corresponding candidate genes are listed in Table 5.

## Discussion

It is known that eggshells are subject to quality loss along with the age process [9]. In egg-type chicken production, there exists a trend to prolong the laying cycles [33], whereas the decline of eggshell quality for old hens proposes a substantial challenge for this development pattern. In the present study, we elucidated genomic architecture underlying the age-dependent dynamic eggshell quality, which has biological and practical significance.

The QTL number detected for eggshell quality was limited in previous studies. The low to medium heritability of the eggshell trait revealed by our results and other study [34] suggested the necessary use of large population for causal mutant identification. Our  $F_2$  population consisting of 1512 hens is the largest population used for eggshell quality GWA analysis so far, and therefore the novel genomic region and loci revealed by the current study should be accurate and reliable. Our GWA analysis for dynamic eggshell quality traits identified thousands of significant associations, that were much more than those screened by previous GWA studies in chicken [10, 35]. These significant mutations were almost in a same region with an extremely strong linkage disequilibrium status, which may result from the insufficient recombination events in the  $F_2$

**Table 4** Contributions of five missense mutations and genomic regions to eggshell quality traits at different wk of age

SNP		<i>rs312347405</i>	<i>rs14837998</i>	<i>rs316793137</i>	<i>rs14840004</i>	<i>rs316607577</i>	Region 1
Chromosome		GGA1	GGA1	GGA1	GGA1	GGA1	GGA1
Position (bp)		64,287,542	65,472,759	67,081,703	67,315,919	67,961,420	57.3–71.4 Mb
Gene symbol		PIK3C2G	RECQL	ABCC9	CASC1	ITPR2	-
EA/AA <sup>a</sup>		C/G	A/G	A/G	T/G	C/T	-
MAF		0.485	0.353	0.414	0.386	0.413	-
Amino acid change		Leu/ Val	Cys/ Arg	Asn/ Asp	Thr/ Pro	Ger/ Gly	-
ESW32	beta (SE) <sup>b</sup>	-0.253 (0.048)	0.276 (0.051)	0.269 (0.049)	-0.151 (0.051)	0.246(0.052)	-
	CPV (%)	5.25	3.36	3.12	0.82	4.90	8.55
ESW40	beta (SE)	-0.361 (0.050)	0.320 (0.050)	0.349 (0.049)	-0.264 (0.050)	0.347(0.051)	-
	CPV (%)	7.21	4.79	5.67	3.08	4.78	12.5
ESW44	beta (SE)	-0.285 (0.046)	0.280 (0.052)	0.302 (0.049)	-0.265 (0.051)	0.291(0.051)	-
	CPV (%)	8.06	3.37	3.97	2.92	6.08	11.17
ESW48	beta (SE)	-0.286 (0.051)	0.228 (0.056)	0.297 (0.053)	-0.195 (0.055)	0.285(0.056)	-
	CPV (%)	7.31	2.33	4.26	1.71	4.93	8.45
ESW52	beta (SE)	-0.321 (0.051)	0.220(0.056)	0.292 (0.053)	-0.179 (0.055)	0.298(0.055)	-
	CPV (%)	7.51	2.21	3.86	1.31	7.77	18.54
ESW56	beta (SE)	-0.333 (0.050)	0.251 (0.053)	0.297 (0.021)	-0.124 (0.053)	0.269(0.053)	-
	CPV (%)	6.8	2.81	4.13	0.52	7.44	15.25
ESW60	beta (SE)	-0.368 (0.054)	0.276 (0.052)	0.295 (0.049)	-0.215 (0.051)	0.291(0.052)	-
	CPV (%)	6.64	3.53	4.11	2.01	5.74	13.77
ESW66	beta (SE)	-0.388 (0.046)	0.252 (0.055)	0.316 (0.052)	-0.234 (0.053)	0.316(0.054)	-
	CPV (%)	8.73	2.86	4.82	2.37	5.86	10.15
ESW72	beta (SE)	-0.382 (0.047)	0.252 (0.056)	-0.332 (0.053)	-0.211 (0.055)	0.292(0.056)	-
	CPV (%)	8.12	2.92	5.19	1.94	6.90	15.14
EST40	beta (SE)	-0.361(0.054)	0.245(0.052)	0.236(0.05)	-0.158(0.051)	0.198(0.053)	-
	CPV (%)	5.33	2.78	2.67	1.07	3.98	5.56
EST44	beta (SE)	-0.313(0.049)	0.211(0.051)	0.225(0.048)	-0.236(0.049)	0.222(0.049)	-
	CPV (%)	4.87	2.01	2.27	2.48	3.42	6.15
EST48	beta (SE)	-0.291(0.053)	0.19(0.054)	0.2(0.052)	-0.141(0.053)	0.199(0.053)	-
	CPV (%)	4.22	1.62	1.86	0.86	4.55	5.44
EST52	beta (SE)	-0.308(0.051)	0.2(0.054)	0.196(0.05)	-0.141(0.052)	0.168(0.052)	-
	CPV (%)	4.63	1.75	1.77	0.81	4.18	5.51
EST56	beta (SE)	-0.312(0.047)	0.222(0.05)	0.234(0.048)	-0.19(0.049)	0.181(0.05)	-
	CPV (%)	4.81	2.21	2.68	1.65	5.33	5.61
EST60	beta (SE)	-0.387(0.051)	0.245(0.053)	0.267(0.05)	-0.25(0.051)	0.212(0.052)	-
	CPV (%)	7.01	2.72	3.42	2.84	4.21	6.92
EST66	beta (SE)	-0.269(0.047)	0.203(0.052)	0.201(0.048)	-0.178(0.049)	0.171(0.05)	-
	CPV (%)	3.49	1.74	1.87	1.33	4.63	4.12
EST72	beta (SE)	-0.339(0.05)	0.165(0.054)	0.237(0.05)	-0.25(0.051)	0.217(0.052)	-
	CPV (%)	5.86	1.21	2.77	3.04	4.02	5.35
ESS32	beta (SE)	-0.253(0.048)	0.172(0.049)	0.129(0.047)	-0.128(0.047)	0.126(0.048)	-
	CPV (%)	3.18	1.28	0.74	0.73	2.65	2.95
ESS40	beta (SE)	-0.361(0.05)	0.245(0.05)	0.264(0.048)	-0.193(0.049)	0.267(0.05)	-
	CPV (%)	6.56	2.80	3.55	1.73	4.82	6.59

**Table 4** Contributions of five missense mutations and genomic regions to eggshell quality traits at different wk of age (*Continued*)

ESS44	beta (SE)	-0.285(0.046)	0.189(0.049)	0.185(0.047)	-0.165(0.048)	0.187(0.048)	-
	CPV (%)	4.04	1.57	1.63	1.21	2.43	4.11
ESS48	beta (SE)	-0.286(0.051)	0.152(0.054)	0.149(0.051)	-0.129(0.052)	0.183(0.052)	-
	CPV (%)	4.16	0.96	1.04	0.78	3.56	4.22
ESS52	beta (SE)	-0.321(0.051)	0.169(0.054)	0.241(0.05)	-0.17(0.052)	0.224(0.052)	-
	CPV (%)	5.30	1.20	2.97	1.39	5.62	5.26
ESS56	beta (SE)	-0.333(0.05)	0.217(0.052)	0.207(0.05)	-0.142(0.051)	0.21(0.052)	-
	CPV (%)	5.64	2.18	2.13	0.92	6.10	5.56
ESS60	beta (SE)	-0.368(0.054)	0.303(0.06)	0.318(0.055)	-0.233(0.058)	0.227(0.058)	-
	CPV (%)	6.70	4.09	4.68	2.29	11.52	16.12
ESS66	beta (SE)	-0.388(0.046)	0.226(0.052)	0.25(0.048)	-0.229(0.049)	0.304(0.049)	-
	CPV (%)	7.59	2.31	3.15	2.43	5.86	6.06
ESS72	beta (SE)	-0.382(0.047)	0.26(0.053)	0.271(0.049)	-0.239(0.051)	0.235(0.051)	-
	CPV (%)	7.35	3.06	3.62	2.70	5.16	6.75

EA effect allele (minor allele), AA alternative allele (major allele), MAF minor allele frequency, CPV contribution to phenotypic variance (%), ESW32, ESW40, ESW44, ESW48, ESW52, ESW56, ESW60 eggshell weights of 32, 40, 44, 48, 52, 56, 60, 66, 72 weeks of age

<sup>a</sup>Estimated allelic substitution effect per copy of the effect allele (EA); SE standard error of the beta <sup>b</sup>Beta means the effect size of minor alleles. Positive/negative effect size value means that the substitution of major allele to minor allele can lead to heavier/lighter yolk or ovary weight

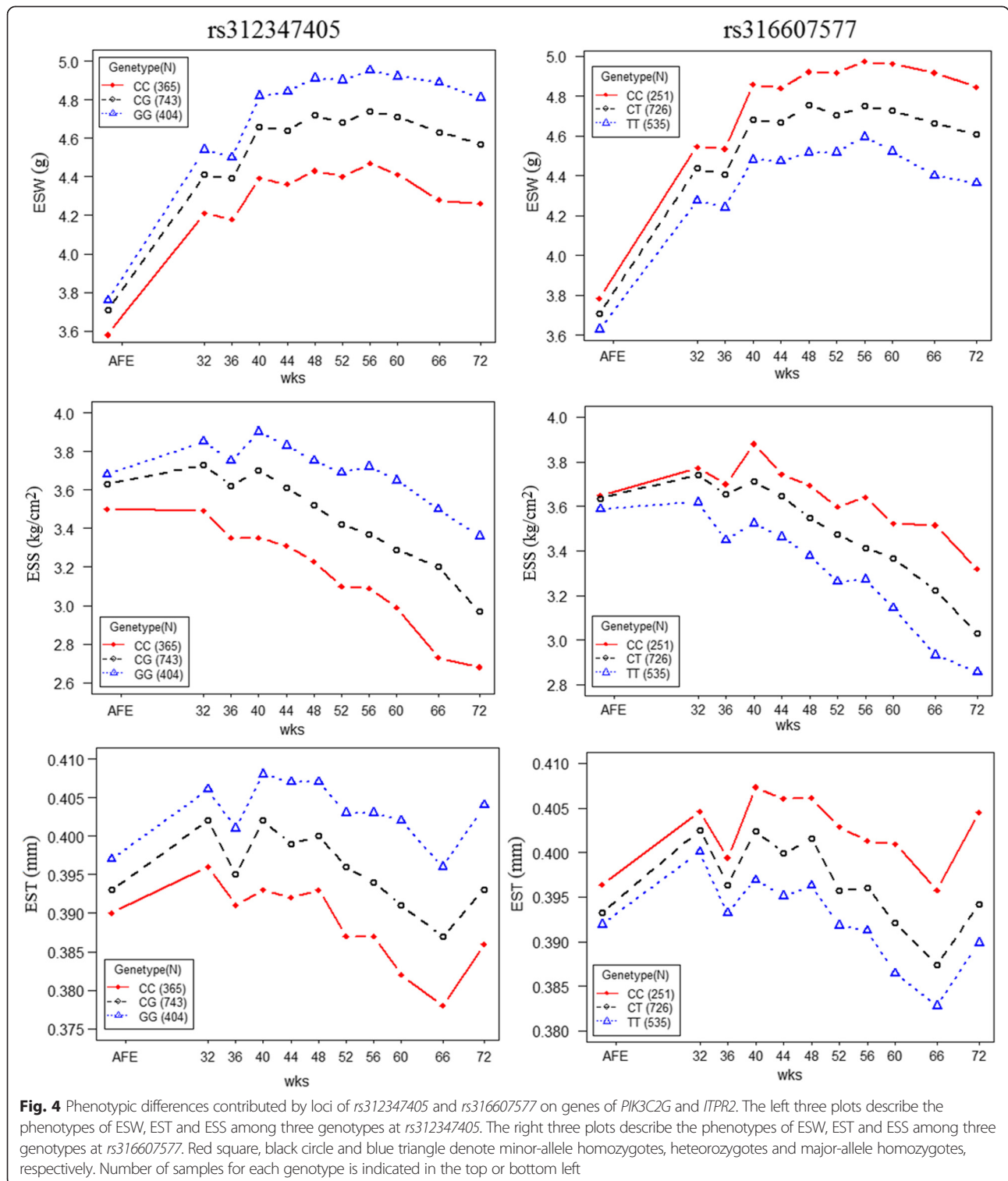
segregation population [36, 37]. Lots of QTL regions affecting ESW, EST and ESS were discovered by previous linkage studies [38] and mainly distributed on GGA 1–9 and GGA11. The promising genomic region from 57.3 to 71.4 Mb identified in the current study had no overlap with the previously reported regions in GGA 1, mainly on 105.2–107.3 Mb, 172.4–173.5 Mb, 180.4–180.5 Mb and 191.5–191.6 Mb (AnimalQTLdb).

The 1531 significant SNPs in this region were distributed on 116 independent genes. We tried to perform Gene Ontology (GO) [39] and pathway analysis [40] on these 116 genes, whereas no significant GO terms or pathways were enriched. This suggested there were not a series of functionally similar genes in this region to influence eggshell quality together. Most SNPs might passively display significant associations for their linkages to a sort of real causal mutants on one or a few genes. Hence, the phenotypic variance explained by all loci in this region was only slightly higher than single locus. It was notable that more associations were discovered for old hens (40–72 weeks) than young hens (AFE to 36 weeks), suggesting certain genetic variants were age-dependent. This was partially consistent with the other GWA studies, indicating no significant associations with both early and late shell quality [10, 41]. However, our results also proved the existence of common genetic variants consistently influencing eggshell quality from 40 to 72 weeks.

Six significant missense mutations in this genomic region were considered as the most important putative variants. After multivariate GWA analysis, only two loci, *rs312347405* and *rs316607577* on *PIK3C2G* and *ITPR2*

in GGA1 remained significantly associated with dynamic eggshell quality traits. The loci of *rs312347405* on *PIK3C2G* could explain the most phenotypic variance of eggshell quality traits among the five missense mutations. Hens with homozygotes of advantageously effective alleles (GG) could produce eggs with excellent ESS, which decreased relatively within a narrow range along with the aging process. *PIK3C2G* encoded proteins belonging to the phosphoinositide 3-kinase (PI3K) family, containing a lipid kinase catalytic domain as well as a C-terminal C2 domain [42], which acted as calcium-dependent phospholipid binding motifs [43]. Previous proteomic screens revealed that a high proportion of lipid-binding proteins abundantly existed in eggshell matrix such as extracellular fatty acid-binding protein (Ex-FABP), prosaposin and apolipoprotein D [14]. Considering the low abundance of lipids in eggshell, the existence of high abundance of lipid-binding proteins in eggshell matrix was perplexed [14]. Furthermore, one association analysis revealed another lipid-related gene, low-density lipoprotein receptor-related protein 8 (LRP8) gene, as a new member of eggshell matrix protein, was significantly associated with eggshell traits [44]. Now in the current study *PIK3C2G* possessing the C2 domain acting as lipid binding motif was also implicated in eggshell property. The role of lipid binding proteins in eggshell or during eggshell formation should be re-recognized. *PIK3C2G* possessing the C2 domain could mediate translocation of proteins to lipid membranes, and also regulate protein-protein interactions in human and mammals [45]. The interoperable matrix proteins and calcite together formed the bioceramic eggshell [46]. It could be hypothesized that lipid binding proteins





might act as carrier or modifier of matrix precursors, which mediate eggshell calcification process or act as frame proteins deposited into eggshell structure. Consequently, *PIK3C2G* was considered as candidate gene for eggshell quality.

The gene of inositol 1,4,5-trisphosphate receptor type 2 (*ITPR2*) harboring the loci of *rs316607577* (exon 25) was the positional and functional candidate gene for eggshell quality revealed by the current GWA analysis. The mutation of *rs316607577* was a nonconservative

**Table 5** Putative genes associated with eggshell quality

Gene symbol <sup>a</sup>	Chr	Tag SNP	Position	Location	AA substitution	SIFT <sup>b</sup>
PIK3C2G	GGA1	rs312347405	64,287,542	Exon 27 of 32	Leu1253Val	0.14
ITPR2	GGA1	rs316607577	67,961,420	Exon 25 of 56	Ser1072Gly	0.40
		rs316447591	67,808,349	3'UTR region	-	-
NCAPG	GGA4	rs14491030	75,486,534	Exon 14 of 21	Val674Ala	0.74

Chr chromosome, AA amino acid, SIFT Sorting intolerant from tolerant

<sup>a</sup>Identification of the gene according to Ensembl genes database 76

<sup>b</sup>SIFT is a program that predicts whether an amino acid substitution affects protein function. Small values means deleterious amino acid change

substitution of serine by glycine (S1072G), with the glycine-encoding allele being associated with stronger eggshell. A gene could be inactivated by a mutation either in a control site or in a coding region [47]. So *rs316447591* on 3'URT of *ITPR2* was also the susceptibility locus. *ITPR2* was known for its mediation in endoplasmic reticulum (ER) calcium release and inositol 1,4,5-trisphosphate (IP3) could mobilize Ca<sup>2+</sup> from intracellular calcium stores to many types of cells [48]. The existence of *ITPR2* in uterine epithelial cell of chicken had been verified [11, 12], furthermore, the expression of *ITPR2* in uterus during eggshell calcification was significantly higher than that in magnum and duodenum which also possess active calcium metabolism [49]. This provided the evidences that *ITPR2* play roles in the regulation of intra-cellular Ca<sup>2+</sup> transportation in uterus and contributed to the process of eggshell calcification. Recently, a genome-wide association study in humans identified *ITPR2* as a susceptibility gene for Kashin-Beck disease which is a chronic osteochondropathy [50], mainly characterized by cartilage degeneration, cartilage matrix degradation, chondrocyte necrosis and apoptosis [51]. This uncovered association with bone disease indicated that it might play a role in general mineralization processes. Birds process medullary bone, a nonstructural type of woven bone, to act as a reservoir for the minerals required for shell calcification [52]. Our previous comparative proteomic analysis for uterine fluid and eggshell matrix proteins also identified that many bone-development related proteins were up-presented in strong eggshell group [15]. Up to date, there only existed one paper referred to the association of ion transporter genes with eggshell quality. They found one SNP on a sodium channel gene (*SCNN1b*) had effects on eggshell strength [53], however, the phenotypic variance of ESS it could explained was relatively low (1.14 %) compared to that of *ITPR2* (2.43–11.52 %) in the current study. All the above evidences suggested *ITPR2* should be a crucial and promising candidate gene relating to eggshell calcification as well as eggshell mechanical property.

Another missense mutation (*rs14491030*) located on non-SMC condensin I complex subunit G (*NCAPG*) gene in GGA4 was discovered in association with eggshell weight for young hens in our GWA analysis. Many other

studies reported multiple genomic regions containing this gene were identified to be associated with body weight [23, 54] and egg weight [35, 55] in chickens and many other growth traits in livestock [56, 57]. The current study suggested that *NCAPG* also involved in influencing eggshell weight. Larger eggs generally owned heavier eggshell. Hence we could not exclude the potential cause that the significant association of *NCAPG* with ESW was just due to its relevance to egg weight. Nevertheless, we still considered *NCAPG* as a candidate gene for its consistent association with eggshell weight from 32 to 40 week.

Previous studies hypothesized that eggshell organic matrix as a complicated mixture of proteins might play a regulatory role in assembly of the calcite zone with calcium carbonate to ultimately determine its strength, and many evidences had been proposed to support this [17], such as the different matrix protein concentrations in strong and weak eggshells [15, 58]. Furthermore, many works conducted to investigate the associations between polymorphisms in genes encoding eggshell matrix proteins and eggshell quality traits [59, 60]. Takahashi *et al.* [60] conducted association analysis between ovocalyxin-32 gene haplotypes and eggshell quality traits in an F<sub>2</sub> intercross and Dunn *et al.* [59] carried out association analysis between polymorphisms in 10 eggshell organic matrix genes and eggshell quality measurements in two lines. These two studies found some significant associations, however, the loci could only explain limited phenotypic variances. Our current study discovered no association between SNPs on eggshell matrix genes and eggshell quality, suggesting that the influences of eggshell matrix genes on eggshell quality were not caused by their nucleotide polymorphisms.

## Conclusions

The dynamic eggshell quality traits at 11 time points from onset of laying to 72 weeks were collected and used for genome-wide association analysis with a 600 K high-density SNP array. According to univariate and multivariate GWA analysis, we discovered a genomic region spanning from 57.3 to 71.4 M in GGA1 significantly associated with eggshell quality. LD and conditional analysis suggested this region were in extremely strong linkage disequilibrium status, especially for the LD block from 64.0 to 67.5 Mb that almost include half (732/1531) of the

total significant loci. Ultimately, three genes, *PIK3C2G*, *ITPR2* and *NCAPG* identified from three missense mutations were considered as promising candidate genes that implicated in dynamic eggshell quality. The homozygotes of advantageously effective alleles on *PIK3C2G* and *ITPR2* possessed the best eggshell quality and could partly counteract the negative effect of aging process. These promising loci and genes could be helpful to engineer practical breeding programs for the improvement of eggshell quality for old hens to meet the need of prolonging the laying cycle. The new findings also have greatly advanced our understanding of the genetic basis underlying eggshell quality.

## Methods

### Resource population

The reciprocal crosses between a standard breed White Leghorn (WL) and a Chinese indigenous strain Dongxiang chickens (DX) were utilized to develop an F<sub>2</sub> resource population. For parents, six WL and six DX males were initially mated with 133 DX and 80 WL females and generated 1029 and 552 chicks for F<sub>1</sub> generation, respectively. Then 25 males and 407 females from WL/DX cross and 24 cocks and 235 hens from DX/WL cross in F<sub>1</sub> generation were used to produce the F<sub>2</sub> population, consisting of 3749 chicks in a single hatch originating from 49 half-sib and 590 full-sib families. The hens were kept in individual cages of the same environment with food and water *ad libitum* at the Jiangsu Institute of Poultry Sciences. Finally, 1534 hens from 49 half-sib families and 550 full-sib families with sufficient phenotypic and pedigree information were selected for SNP genotyping.

### Phenotypic measurements

Eggshell quality traits including eggshell weight (ESW), eggshell thickness (EST) and eggshell strength (ESS) were measured for the first egg of hens and then at 32, 36, 40, 44, 48, 52, 56, 60, 66 and 72 weeks of age, consisting of 11 time points. Except for the first egg, fresh eggs were collected for 4 successive days to ensure two eggs per hen. Then breaking strength (pole to pole) of each egg was measured vertically with the Eggshell Force Gauge (Model-II, Robotmation, Tokyo, Japan). After that, eggs were broken to remove the internal contents and the remaining eggshells were washed clean with tap water. After drying in the air at room temperature, eggshell weights and eggshell thickness were measured with electronic scale and eggshell thickness gauge (FHK, Tokyo, Japan) respectively. Descriptive statistics were calculated with R project (R version 3.1.2) using all available records. The function of 'rnttransform' in the GenABEL package of R project was used for the rank-based inverse normal transformations (INTs) [61].

### Genotyping and quality control (QC)

Genomic DNA was extracted by standard phenol/chloroform method and genotyped with the 600 K Affymetrix Axiom Chicken Genotyping Array (Affymetrix, Inc. Santa Clara, CA, USA). Affymetrix Power Tools v1.16.0 (APT) software was then used for quality control and genotype calling. Specifically, only samples with dish quality control (DQC) > 0.82 and call rate > 97 % were included into the subsequent analyses. An R script supplied by Affymetrix was run to compute the SNP QC metrics and filter out individual SNPs falling below given thresholds. After these QC steps, 1512 individuals and 532,299 SNPs remained valid. Given the current statistical methods were more powerful for the detection of the associations between phenotypes and autosomal genotypes, all SNPs on sex chromosomes were excluded in QC process. The PLINK v1.90 package [62] were then used for further quality control to improve the detect power, in which SNPs with minor allele frequency (MAF) < 5 % and Hardy-Weinberg equilibrium (HWE) test  $P < 1 \times 10^{-6}$  were removed from downstream analysis. Some sporadic missing genotypes were imputed using the BEAGLE v4.0 procedure [63], only SNPs with imputation quality score  $R^2 > 0.5$  were retained. Ultimately, up to 1512 individuals and 435,867 SNPs were valid for the following GWA analysis.

### Genome-wide Association (GWA) analysis

Principal component analysis (PCA) implemented in PLINK package was first conducted prior to GWA analysis to eliminate spurious associations resulting from the presence of cryptic relatedness or hidden population stratification. To properly decide the thresholds for genome-wide significant/suggestive associations, a method of simpleM [64] was used to corrected the number of multiple tests. With this approach, an effective number of 59,308 independent tests was suggested, hence the genome-wide significant and suggestive  $P$ -values were  $8.43 \times 10^{-7}$  (0.05/59,308) and  $1.69 \times 10^{-5}$  (1.00/59,308), respectively.

Univariate association test equipped with an exact mixed model approach in GEMMA v0.94 software [65] was first performed with the valid individuals and SNPs. The centered relatedness matrix was calculated by those independent SNPs, and then the derived Wald test  $P$ -value was calculated for the significance level between SNPs and phenotypes. The univariate linear mixed model is as follows:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{x}\boldsymbol{\beta} + \mathbf{u} + \boldsymbol{\varepsilon}$$

$\mathbf{y}$  is an  $n \times 1$  vector of phenotypic values for  $n$  individuals,  $\mathbf{W}$  is an  $n \times c$  matrix of covariates (fixed effects, *i.e.*, top five PCs) including a column vector of

1,  $\alpha$  is a  $c \times 1$  vector of corresponding coefficients including the intercept,  $\mathbf{x}$  is an  $n \times 1$  vector of marker genotypes at the locus tested,  $\beta$  is the corresponding effect size of the marker and all effects are reported for the minor allele in each marker,  $\mathbf{u}$  is an  $n \times 1$  vector of random polygenic effects with a covariance structure as  $\mathbf{u} \sim N(0, \mathbf{K}V_g)$ , where  $\mathbf{K}$  represents a known  $n \times n$  genetic relatedness matrix derived from SNP markers and  $V_g$  is the polygenic additive variance, and  $\varepsilon$  is an  $n \times 1$  vector of random residuals with  $\varepsilon \sim N(0, \mathbf{I}V_e)$ , where  $\mathbf{I}$  is an  $n \times n$  identity matrix, and  $V_e$  is the residual variance component. We used the Wald test statistic  $F_{\text{Wald}} = \hat{\beta}^2 / \text{Var}(\hat{\beta})$  for each SNP to test the null hypothesis  $\beta = 0$ , where the best linear unbiased estimate (BLUE) of  $\beta$  and the corresponding sampling variance  $\text{Var}(\hat{\beta})$  are obtained by solving the mixed model equations (MME) based on estimated  $V_g$  and  $V_e$ .

The manhattan plots and quantile-quantile (QQ) plots were drawn by the “gap” packages [66] in R project. The genomic inflation factor  $\lambda$  which were the judgements for the extent of false positive signals was also calculated with the function of `estlambda` implemented in the GenABEL package [61] in R project.

To further analyze genetic variants affecting the dynamic eggshell quality traits along with aging process, a multivariate association analysis [67] that directly utilized phenotypes from multiple time points on an individual was applied. Only the time points and SNPs that had significant associations in univariate model were also included in this multivariate model. For each SNP marker, a multivariate linear mixed model could be fitted in the following form:

$$\mathbf{Y} = \mathbf{W}\mathbf{A} + \mathbf{x}\beta^T + \mathbf{G} + \mathbf{E}$$

where  $\mathbf{Y}$  is an  $n$  by  $d$  matrix of  $d$  phenotypes for  $n$  individuals,  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_c)$  is an  $n \times c$  matrix of covariates (fixed effects, *i.e.*, top five PCs) including a column of 1s,  $\mathbf{A}$  is a  $c$  by  $d$  matrix of corresponding coefficients including the intercept,  $\mathbf{x}$  is an  $n$ -vector of marker genotypes,  $\beta$  is a  $d$  vector of marker effect sizes for the  $d$  phenotypes. It should be noted that  $\mathbf{G}$  is an  $n$  by  $d$  matrix of random effects with  $\mathbf{G} \sim \text{MVN}_{n \times d}(0, \mathbf{K}, \mathbf{V}_g) \sim \text{MVN}_{n \times d}(0, \mathbf{K}, \mathbf{V}_g)$  where  $\mathbf{V}_g$  is a  $d$  by  $d$  symmetric positive definite matrix of genetic variance component, and  $\mathbf{E}$  is an  $n$  by  $d$  matrix of residual errors with  $\mathbf{E} \sim \text{MVN}_{n \times d}(0, \mathbf{I}, \mathbf{V}_e)$  where  $\mathbf{V}_e$  is a  $d$  by  $d$  symmetric positive definite matrix of residual variance component ( $\mathbf{K}$  and  $\mathbf{I}$  are the same in the two models).  $\text{MVN}_{n \times d}(0, \mathbf{V}_1, \mathbf{V}_2)$  denotes the  $n \times d$  matrix normal distribution with mean 0, row covariance matrix  $\mathbf{V}_1$  ( $n$  by  $n$ ) and column covariance matrix  $\mathbf{V}_2$  ( $d$  by  $d$ ).

### Conditional and Linkage disequilibrium (LD) analysis

Notably, many SNPs maybe passively significant associated with target traits, resulting from their strong linkage to a really causal mutants. To demarcate independent association signals in a putative region, the conditional analyses as well as the LD analysis were both performed in multivariate and univariate models, through fitting the genotypes of one candidate signal as covariates [62]. In general, GWAS does not distinguish a genuine causal locus from those statistically significant loci within a strong LD region. Therefore, in order to characterize potential candidate genes responsible for a trait, we firstly conducted a LD analysis and inferred the haplotype blocks containing peak SNPs by Haploview v4.2 [68]. A block is derived using the solid spine algorithm, and defined as that the first and last SNPs in a region are in strong LD ( $D' \geq 0.8$ ) with all intermediate SNPs.

### Estimation of heritability and contribution to phenotypic variance (CPV)

Univariate restricted maximum likelihood (REML) implemented in GCTA v1.24 program [69] was performed to estimate the heritability explained by the eligible SNPs ( $h_{\text{SNP}}^2$ ). We also quantified the pair-wise genetic and phenotypic correlations for each trait at multiple time points with the bivariate mixed model [70]. A genetic relationship matrix (GRM) derived from all genotyped SNPs on autosomes and two linkage groups was created, and the top five PCs calculated by the GCTA tool were included as covariates to account for the potential population structure. We then estimated the phenotypic variance contributed by these associated loci or genomic region.

### Gene identification

Variant Effect Predictor (VEP) and Biomart tools based on the Galgal4 assembly supported by Ensembl were used for the identification of candidate genes that the significant loci located on [71] and detection the genes within in a given genomic region [72].

### Additional files

**Additional file 1: Figure S1.** Change curve of eggshell weight (ESW), eggshell percentage (ESP), eggshell thickness (EST) and eggshell strength (ESS) along with the age of laying hens. Plots A, B, C, D displayed the change curve of ESW, ESP, EST and ESS respectively. **Figure S2.** Manhattan plot (left) and quantile-quantile plot (right) of the observed  $P$ -values for ESW, EST and ESS at age of first egg and at 32, 36, 40, 48, 52, 56, 60, 66, 72 weeks of old. The Manhattan plot indicates  $-\log_{10}$  (observed  $P$ -values) for genome-wide SNPs ( $y$ -axis) plotted against their respective positions on each chromosome ( $x$ -axis), and the horizontal green and black lines depict the genome-wide suggestive ( $1.69 \times 10^{-5}$ ) and significant ( $8.43 \times 10^{-7}$ ) threshold, respectively. For quantile-quantile plot, the  $x$ -axis shows the expected  $-\log_{10}$ -transformed  $P$ -values, and the  $y$ -axis represents the observed  $-\log_{10}$ -transformed  $P$ -values. The genomic inflation factors ( $\lambda$ ) are shown on the top left in the QQ plot. Green points represent the genome-wide significant associations. **Figure S3.** Regional plots and

conditional analysis in multivariate model for eggshell thickness (EST) and eggshell strength (ESS). Plot A: regional and conditional plot for EST. Plot B: regional and conditional plot for ESS. (PDF 4277 kb)

**Additional file 2: Table S1. Summary of genetic analysis for eggshell thickness at different wks of age. Table S2. Summary of genetic analysis for eggshell strength at different wks of age. Table S3.** Genome-wide significant SNPs for eggshell weight (ESW), eggshell thickness (EST) and eggshell strength (ESS) at different age by univariate model. **Table S4.** Genome-wide significant SNPs for eggshell weight (ESW), eggshell thickness (EST) and eggshell strength (ESS) by multivariate model. (XLSX 1372 kb). **Table S5.** Significant loci that located on 3' or 5' UTRs.

#### Abbreviations

GWAS: Genome-wide association study; GWA: Genome-wide association; SNP: Single nucleotide polymorphism; QTL: Quantitative trait loci; LD: Linkage disequilibrium; REML: Restricted maximum likelihood; GRM: Genetic relationship matrix; MAF: Minor allele frequency; HWE: Hardy-Weinberg equilibrium; PCA: Principal component analysis; VEP: Variant effect predictor.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

NY and KW conceived the study and designed the project. CS, LQ, GY and JY performed genetic and bioinformatics analyses. MS, LJQ, ZD, GX contributed to collecting samples and measuring phenotypic data. CS, LQ wrote and NY revised the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

The current research was funded in part by National Natural Science Foundation of China (31472084), National High Technology Development Plan of China (2013AA102501), Programs for Changjiang Scholars and Innovative Research in University (IRT1191), and China Agriculture Research Systems (CARS-41).

#### Author details

<sup>1</sup>National Engineering Laboratory for Animal Breeding and MOA Key Laboratory of Animal Genetics and Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China. <sup>2</sup>Jiangsu Institute of Poultry Science, Yangzhou, Jiangsu 225125, China.

Received: 14 January 2015 Accepted: 22 July 2015

Published online: 31 July 2015

#### References

- Chien YC, Hincke MT, McKee MD. Ultrastructure of avian eggshell during resorption following egg fertilization. *J Struct Biol*. 2009;168(3):527–38.
- Nys Y, Gautron J, Garcia-Ruiz JM, Hincke MT. Avian eggshell mineralization: biochemical and functional characterization of matrix proteins. *Comptes Rendus Palevol*. 2004;3(6–7):549–62.
- Nys Y, Bain M, Van Immerseel F. Improving the safety and quality of eggs and egg products. Cambridge, Woodhead Publishing Limited; 2011.
- Zeidler G: Processing and packaging shell eggs. In: Commercial Chicken Meat and Egg Production. Kluwer Academic Publishers Springer US; 2002: 1129–1161.
- Tuiskula-Haavisto M, Honkatukia M, Preisinger R, Schmutz M, de Koning DJ, Wei WH, et al. Quantitative trait loci affecting eggshell traits in an F(2) population. *Anim Genet*. 2011;42(3):293–9.
- Lourens A, Van den Brand H, Meijerhof R, Kemp B. Effect of eggshell temperature during incubation on embryo development, hatchability, and posthatch development. *Poultry Sci*. 2005;84(6):914–20.
- Deeming DC, Ferguson MW. Egg incubation: its effects on embryonic development in birds and reptiles. Cambridge: Cambridge University Press; 1991.
- Meir M, Ar A. Changes in eggshell conductance, water loss and hatchability of layer hens with flock age and moulting. *Brit Poultry Sci*. 2008;49(6):677–84.
- Al-Batshan H, Scheideler S, Black B, Garlich J, Anderson K. Duodenal calcium uptake, femur ash, and eggshell quality decline with age and increase following molt. *Poultry Sci*. 1994;73(10):1590–6.
- Liu W, Li D, Liu J, Chen S, Qu L, Zheng J, et al. A genome-wide SNP scan reveals novel loci for egg production and quality traits in white leghorn and brown-egg dwarf layers. *PLoS One*. 2011;6(12):1427–30.
- Jonchere V, Rehault-Godbert S, Hennequet-Antier C, Cabau C, Sibut V, Cogburn LA, et al. Gene expression profiling to identify eggshell proteins involved in physical defense of the chicken egg. *BMC Genomics*. 2010;11:57.
- Brionne A, Nys Y, Hennequet-Antier C, Gautron J. Hen uterine gene expression profiling during eggshell formation reveals putative proteins involved in the supply of minerals or in the shell mineralization process. *BMC Genomics*. 2014;15(13):1736–42.
- Liu Z, Zheng Q, Zhang X, Lu L. Microarray analysis of genes involved with shell strength in layer shell gland at the early stage of active calcification. *Asian Austral J Anim*. 2013;26(5):609–24.
- Mann K, Macek B, Olsen JV. Proteomic analysis of the acid-soluble organic matrix of the chicken calcified eggshell layer. *Proteomics*. 2006;6(13):3801–10.
- Sun C, Xu G, Yang N. Differential label-free quantitative proteomic analysis of avian eggshell matrix and uterine fluid proteins associated with eggshell mechanical property. *Proteomics*. 2013;13(23–24):3523–36.
- Rose-Martel M, Du J, Hincke MT. Proteomic analysis provides new insight into the chicken eggshell cuticle. *J Proteome*. 2012;75(9):2697–706.
- Marie P, Labas V, Brionne A, Harichaux G, Hennequet-Antier C, Nys Y, et al. Quantitative proteomics and bioinformatic analysis provide new insight into protein function during avian eggshell biomineralization. *J Proteome*. 2015;113:178–93.
- Rose-Martel M, Smiley S, Hincke MT. Novel identification of matrix proteins involved in calcitic biomineralization. *J Proteome*. 2015;116:81–96.
- Panheleux M, Bain M, Fernandez MS, Morales I, Gautron J, Arias JL, et al. Organic matrix composition and ultrastructure of eggshell: a comparative study. *Brit Poultry Sci*. 1999;40(2):240–52.
- Rodriguez-Navarro AB. Rapid quantification of avian eggshell microstructure and crystallographic-texture using two-dimensional X-ray diffraction. *Brit Poultry Sci*. 2007;48(2):133–44.
- Riley A, Sturrock CJ, Mooney SJ, Luck MR. Quantification of eggshell microstructure using X-ray micro computed tomography. *Brit Poultry Sci*. 2014;55(3):311–20.
- Wardecka B, Olszewski R, Jaszczak K, Zieba G, Pierzchala M, Wicinska K. Relationship between microsatellite marker alleles on chromosomes 1–5 originating from the Rhode Island Red and Green-legged Partridge breeds and egg production and quality traits in F(2) mapping population. *J Appl Genet*. 2002;43(3):319–29.
- Sasaki O, Odawara S, Takahashi H, Nirasawa K, Oyamada Y, Yamamoto R, et al. Genetic mapping of quantitative trait loci affecting body weight, egg character and egg production in F2 intercross chickens. *Anim Genet*. 2004;35(3):188–94.
- Wright D, Kerje S, Lundstrom K, Babol J, Schutz K, Jensen P, et al. Quantitative trait loci analysis of egg and meat production traits in a red junglefowl x White Leghorn cross. *Anim Genet*. 2006;37(6):529–34.
- Ball AD, Stapley J, Dawson DA, Birkhead TR, Burke T, Slate J. A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (*Taeniopygia guttata*). *BMC Genomics*. 2010;11:218.
- Fernandez ME, Goszczynski DE, Liron JP, Villegas-Castagnasso EE, Carino MH, Ripoli MV, et al. Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genet Mol Biol*. 2013;36(2):185–91.
- Yu H, Xie W, Wang J, Xing Y, Xu C, Li X, et al. Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS One*. 2011;6(3):e17595.
- Kranis A, Gheyas AA, Boschiero C, Turner F, Yu L, Smith S, et al. Development of a high density 600 K SNP genotyping array for chicken. *BMC Genomics*. 2013;59(5):670–70.
- Koch RM, Clark R. Influence of sex, season of birth and age of dam on economic traits in range beef cattle. *J Anim Sci*. 1955;14(2):386–97.
- van de Pol M, Verhulst S. Age-dependent traits: a new statistical model to separate within-and between-individual effects. *Am Nat*. 2006;167(5):766–73.
- Costanza MC, Beer-Borst S, James RW, Gaspoz J-M, Morabia A. Consistency between cross-sectional and longitudinal SNP: blood lipid associations. *Eur J Epidemiol*. 2012;27(2):131–8.

32. Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beatty TH, et al. Genetic epidemiology of COPD (COPDGene) study design. *COPD*. 2011;7(1):32–43.
33. Schulte-Drüggel R, Thiele H. Determining the optimum replacement schedule for commercial layers: does molting pay off? *Small*. 2013;42(50):658.
34. Zhang L-C, Ning Z-H, Xu G-Y, Hou Z-C, Yang N. Heritabilities and genetic and phenotypic correlations of egg quality traits in brown-egg dwarf layers. *Poultry Sci*. 2005;84(8):1209–13.
35. Wolc A, Arango J, Jankowski T, Dunn I, Settar P, Fulton JE, et al. Genome-wide association study for egg production and quality in layer chickens. *J Anim Breed Genet*. 2014;131(3):173–82.
36. Li N, Stephens M. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*. 2003;165(4):2213–33.
37. Begovich A, McClure G, Suraj V, Helmuth R, Fildes N, Bugawan T, et al. Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J Immunol*. 1992;148(1):249–58.
38. Hu Z-L, Fritz ER, Reecy JM. AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. *Nucleic Acids Res*. 2007;35 suppl 1:D604–9.
39. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000;25(1):25–9.
40. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*. 2011;40(Database issue):D109–14. doi:10.1093/nar/gkr988. Epub 2011 Nov 10.
41. Abasht B, Sandford E, Arango J, Settar P, Fulton JE, O'Sullivan NP, et al. Extent and consistency of linkage disequilibrium and identification of DNA markers for production and egg quality traits in commercial layer chicken populations. *BMC Genomics*. 2009;10 Suppl 2:S2.
42. Walker EH, Perisic O, Ried C, Stephens L, Williams RL. Structural insights into phosphoinositide 3-kinase catalysis and signalling. *Nature*. 1999;402(6759):313–20.
43. Welters P, Takegawa K, Emr SD, Chrispeels MJ. AtVPS34, a phosphatidylinositol 3-kinase of *Arabidopsis thaliana*, is an essential protein with homology to a calcium-dependent lipid binding domain. *Proc Natl Acad Sci*. 1994;91(24):11398–402.
44. Yao JF, Chen ZX, Xu GY, Wang XL, Ning ZH, Zheng JX, et al. Low-density lipoprotein receptor-related protein 8 gene association with egg traits in dwarf chickens. *Poult Sci*. 2010;89(5):883–6.
45. Balla T. Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci*. 2005;118(10):2093–104.
46. LEACH RM. Biochemistry of the organic matrix of the eggshell. *Poultry Sci*. 1982;61(10):2040–7.
47. Lewin B. GENES VIII Part 1.25 Proteins are trans-acting but sites on DNA are cis-acting. Pearson Prentice Hall: Pearson Education, Inc. Upper Saddle River, NT 07458; 2004.
48. Goto J, Mikoshiba K. Inositol 1,4,5-trisphosphate receptor-mediated calcium release in Purkinje cells: from molecular mechanism to behavior. *Cerebellum*. 2011;10(4):820–33.
49. Jonchere V, Brionne A, Gautron J, Nys Y. Identification of uterine ion transporters for mineralisation precursors of the avian eggshell. *BMC Physiol*. 2012;12:10.
50. Zhang F, Wen Y, Guo X, Zhang Y, Wang X, Yang T, et al. Genome-wide association study identifies ITPR2 as a susceptibility gene for Kashin-Beck disease in Han Chinese. *Arthritis Rheumatol*. 2015;67(1):176–81.
51. Hinsenkamp M. Kashin-Beck disease. *Int Orthop*. 2001;25(3):133–3.
52. Dacke C, Arkle S, Cook D, Wormstone I, Jones S, Zaidi M, et al. Medullary bone and avian calcium regulation. *J Exp Biol*. 1993;184(1):63–88.
53. Fan YF, Hou ZC, Yi GQ, Xu GY, Yang N. The sodium channel gene family is specifically expressed in hen uterus and associated with eggshell quality traits. *BMC Genet*. 2013;14:90.
54. Gu X, Feng C, Ma L, Song C, Wang Y, Da Y, et al. Genome-wide association study of body weight in chicken F2 resource population. *PLoS One*. 2011;6(7):e21872.
55. Wolc A, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, et al. Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. *Anim Genet*. 2012;43 Suppl 1:87–96.
56. Eberlein A, Takasuga A, Setoguchi K, Pfuhl R, Flisikowski K, Fries R, et al. Dissection of genetic factors modulating fetal growth in cattle indicates a substantial role of the non-SMC condensin I complex, subunit G (NCAPG) gene. *Genetics*. 2009;183(3):951–64.
57. Lindholm-Perry AK, Sexten AK, Kuehn LA, Smith TP, King DA, Shackelford SD, et al. Association, effects and validation of polymorphisms within the NCAPG - LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. *BMC Genet*. 2011;12:103.
58. Ahmed A, Rodriguez-Navarro A, Vidal M, Gautron J, Garcia-Ruiz JM, Nys Y. Changes in eggshell mechanical properties, crystallographic texture and in matrix proteins induced by moult in hens. *Brit Poultry Sci*. 2005;46(3):268–79.
59. Dunn IC, Joseph NT, Bain M, Edmond A, Wilson PW, Milona P, et al. Polymorphisms in eggshell organic matrix genes are associated with eggshell quality measurements in pedigree Rhode Island Red hens. *Anim Genet*. 2009;40(1):110–4.
60. Takahashi H, Sasaki O, Nirasawa K, Furukawa T. Association between ovocalyxin-32 gene haplotypes and eggshell quality traits in an F2 intercross between two chicken lines divergently selected for eggshell strength. *Anim Genet*. 2010;41(5):541–4.
61. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics*. 2007;23(10):1294–6.
62. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
63. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet*. 2009;84(2):210–23.
64. Gao X, Becker LC, Becker DM, Starmer JD, Province MA. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet Epidemiol*. 2010;34(1):100–5.
65. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44(7):821–4.
66. Zhao JH: gap: Genetic Analysis Package. R package version 1.1–12. <http://cran.r-project.org/web/packages/gap/index.html> 2014.
67. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat Methods*. 2014;11(4):407–9.
68. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263–5.
69. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76–82.
70. Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet*. 2012;44(3):247–50.
71. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010;26(16):2069–70.
72. Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudis G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)*. 2011;2011(1):56–65.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

