TÜBİTAK

Turkish Journal of Veterinary and Animal Sciences

http://journals.tubitak.gov.tr/veterinary/

Research Article

Turk J Vet Anim Sci
(2020) 44: 9-16

doi:10.3906/vet-1904-33

Functional and whole regression-based genome-wide association analyses for weight measurements of chicken eggs

BURAK KARACAÖREN*

Department of Animal Science, Faculty of Agriculture, Akdeniz University, Antalya, Turkey

Received: 11.04.2019 • **Accepted/Published Online:** 14.12.2019 • **Final Version:** 10.02.2020

Abstract: In genomic studies, complex traits can be modelled using repeated measures, thereby gaining a better understanding of the underlying biology. An increased number of measurements per individual might reduce measurement noise, thus increasing the likelihood of detecting true genomic signals. Here we aimed to predict genomic signals over a logistic curve referring to multiple underlying genetic architectures, for both simulated and longitudinal egg weight datasets. The chicken data were obtained from 92 sires and 801 dams of an 11th generation pure line, resulting in data from 1078 hens. We analysed longitudinal measurements of egg weights with 294,705 single nucleotide polymorphisms (SNPs). We found that a single regression-based functional genome-wide association study (fGWAS) could be useful for manipulating dynamic egg weight over the entire laying period based on a moderate to major effect gene. The fGWAS SNPs associated with the egg weight were located on chromosome 1 close to the gene *DLEU7*, which has a role in regulating ovary weight in chickens. The SNPs were detected based on the absolute effect sizes using whole regression Bayesian models. This approach is likely to be useful for predicting polygenic risk scores and/or genomic breeding values during the genomic selection/gene editing for longitudinal egg weight measurements.

Key words: Egg weight, genomic selection, Gibbs sampling, genome-wide association mapping

1. Introduction

In genomic studies, complex traits can be modelled using repeated measures to achieve a better understanding of the underlying biology [1,2]. An increasing number of available measurements per individual might reduce the measurement noise and thus increase the detection of true genomic signals [3,4]. However, cross-sectional genome-wide association studies (GWASs) are commonly used to investigate correlations between single nucleotide polymorphisms (SNPs) and complex phenotypes. In genomic studies, abstracting a dynamic biological process of complex phenotypes into a single observation (i.e. averaging of observations) [5,6] or discarding the correlations among the available measurements over time might lead to suboptimal conclusions [7,8].

Using single SNP regression methodologies, GWASs are used to identify sources of variations. The problem of whole genetic variation in complex phenotypes cannot be accounted for by associated SNPs, termed "missing heritability" [9]. Two seemingly divergent assumptions have been applied in the development of GWAS models: searching for major genes using single SNP regression models and assuming that at least one SNP is correlated with all genes. This leads to the simultaneous use of whole SNPs [10] with the underlying genetic architecture of the phenotype. Compared to a single regression approach, employing whole genome regressions can better explain genomic variation [11] for some phenotypes in some organisms.

In chicken farming, longitudinal measures of egg weight (EW) are commonly recorded for breeding purposes. EW is influenced by both environmental and genetic factors. Different from other organisms, livestock (particularly chicken) experiments have pruned the sources of variations due to the use of homogenized lines and controlled environmental factors [12]. Based on the above discussion, to identify the genomic locations associated with EW, it is necessary to consider assumptions about the genetic architecture of the EW with special reference to the longitudinal nature [13] of the measurements. Recently, Liu et al. [14] detected time specific genomic signals for EW using separate association tests. The main objective of this study is to predict the genomic signals over a logistic curve referring to multiple underlying genetic architectures, for both simulated [15] and longitudinal EW [14] datasets.

* Correspondence: burakkaracaoren@akdeniz.edu.tr

2. Materials and methods

2.1. Materials

The pedigree included 100 full-sib families, each with 20 progenies. The longitudinal quantitative phenotypes were simulated at five time points (0, 132, 265, 397, and 530 days). The genome consisted of 453 SNPs distributed over 5 chromosomes and across 5 Morgan (M). The details of the simulated dataset can be found in [15].

The chicken data were obtained from 92 sires and 801 dams of an 11th generation Rhode Island pure line, with a total of 1078 hens [14]. We analysed longitudinal measurements of the initial EW, and again at 28, 36, 56, 66, and 72 weeks old. Before statistical analyses, SNPs were removed from the data if the Hardy–Weinberg equilibrium P-value was less than 1e-6, or if the minor allele frequency was less than 1%. Finally, a total of 294,705 SNPs were collected for the genomic analyses. The details of the chicken dataset can be found in [14].

2.2. Methods

We applied various genomic association models to explore the different genetic architectures of the longitudinal phenotypes in both simulated QTLMAS and the chicken dataset. Here we used the logistic growth function to model the repeated measurements. The logistic growth curve was considered the most common model to define the time specific dynamics of the biological process [16,17]:

$$y(t) = \frac{a}{1 + (b-1)e^{-rt}} \quad (1)$$

where y(t) is the measurement at time t, a is the asymptote of the measurement of the animal, r is the steepness (growth rate) of the curve, and b is the ratio of current and initial measurements.

We used GRAMMAR and EGSCORE (conducting genome-wide rapid association using mixed model and regression) functions in the GenABEL package [18] for single SNP regression association analyses to estimate the parameters of the logistic growth curve (Eq. (1)). The GRAMMAR (raw, genomic control, or gamma versions) function corrects the phenotypes using a genomic relationship matrix by:

$$y = \mathbf{X}d + \mathbf{Z}a + e \quad (2)$$

where y contains the observations (estimated logistics curve parameters as a, b, and r), d is the fixed effects, a is the additive genetic effect, matrices **X** and **Z** are the incidence matrices, and **e** is a vector containing residuals.

$$Var\begin{pmatrix} a \\ e \end{pmatrix} N\left[\mathbf{0}; \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix}\right]$$

For the random effects, it is assumed that **A** is the additive genomic relationship matrix for the animals, **I** is an identity matrix, σ^2_a is the additive genetic variance, and σ^2_e is the residual variance. In the second step, the association could be detected by:

$$y = \mathbf{X}f + \boldsymbol{\eta} + \mathbf{e} \quad (3)$$

where **y** represents a vector of residuals obtained from Eq. (2), **η** is an intercept, **X** is a design matrix relating observations with **f** regression coefficients vector to be estimated, and **e** is a vector of residuals assumed to be normally distributed. Population stratification could also be corrected by the principal components using the EGSCORE [19] function implemented in GenABEL. We also used the functional genome-wide association (fGWAS) model to combine longitudinal phenotypes to increase statistical power for detecting genomic signals [3,6]. In fGWAS, the time dependent mean value of genotype j at time t could be used to predict the additive effect of the SNP, a(t):

$$a(t) = \frac{1}{2}\left[\mu_{1(t)} - \mu_{3(t)}\right]$$

where $\mu_{1(t)}$ and $\mu_{3(t)}$ represents *AA* and *aa* homozygote genotypes at time t, respectively. Due to the huge number of multiple hypotheses for testing, we used a false discovery rate approach to reduce type 1 errors in single SNP regression models.

We used the Bayesian lasso (BL) and Bayes C(π) models [20] to predict SNP effects using Gibbs sampling [21]:

$$y_i = sex + \sum_{j=1}^{n}\left(z_{ij}a_j\delta_j\right) + e_i$$

where y_i is the phenotype of the ith animal for the estimated parameters of the logistic curve from Eq.(1); z_{ij} is an indicator variable for the ith animal, jth SNP locus, and kth allele; a_j is the marker locus effect; δ_j indicates if the SNP has an effect or not; and e_i is the residual for animal i. For each phenotype, the Markov chain Monte Carlo (MCMC) algorithm was run for 1,000,000 samples, and the first 2000 samples were discarded as the burn-in period. As a thinning method, we collected each 10th sample from each realization of the MCMC. We used the default value of GS3 so that the proportion of quantitative trait loci (QTLs) (π) in the whole genome was 0.01. For comparison purposes, we standardized the SNP effects by subtracting the mean and dividing by the standard deviation of the SNPs.

3. Results

GWASs are commonly used to locate associated genetic loci with quantitative phenotypes in domestic animals. Using different statistical models, we investigated the effects of various assumptions about the genetic architecture of EW in chickens and QTLMAS datasets.

3.1. QTLMAS results

We used a logistic growth model [22] to detect genomic signals in the QTLMAS dataset. The GWAS for the QTLMAS dataset was performed using each of the

logistic growth model parameters. The estimated genomic heritabilities were 0.061, 0.234, and 0.153 for the asymptote (A), inflection point (I), and scaling factor (S) parameters, respectively. The existence of true genomic signals was inferred by comparing our predicted SNP locations with QTLMAS organizer results at 10 cM intervals. Different from single SNP regression models, Bayesian models simultaneously use whole markers in association with the phenotype (achieved through different prior distributions, and hence different genetic architectures). The predicted SNP effects were sorted based on the absolute standardized regression effects (with minimum posterior inclusion to the model of 0.001) for the Bayesian models. The mean heterozygosis (and its standard error) of SNPs and animals were 0.2072 (0.1618) and 0.2126 (0.028), respectively.

We hypothesized various genetic architectures for the simulated QTLMAS dataset. The single SNP regression models were used to search for major genes by the genomic relationship matrix (GRAMMAR) and principal component (EGSCORE) approaches. The results of the single SNPs and Bayesian analyses are summarized in Table 1. Genome-wide significant SNPs were identified for all logistics curve parameters. By carrying out a single SNP regression GWAS, we detected up to 13 SNPs in GRAMMAR and 18 SNPs in PCR approaches using a corrected threshold for values of $P < 0.05$. Genomic inflation factors were predicted to be between 1 and 1.296 by GRAMMAR and 2.001 and 4.464 by EGSCORE approaches. By Bayesian models, we detected up to 24 SNPs in the Bayes C(π) model and 12 in the BL model.

Table 1. The number of true/total significant SNPs (number of true QTLs) from GWA obtained from the estimated parameters of logistics curve (asymptote, inflection point, and scaling factor) using GRAMMAR, EGSCORE, Bayes C(π), and Bayesian lasso (BL) models.

	Asymptote	Inflection point	Scaling factor
GRAMMAR	6/14	5/14	2/16
EGSCORE	6/20	9/16	3/20
Bayes C(π)	10/20	8/20	6/20
BL	4/20	4/20	4/20

Unlike the 2-step approaches, the fGWAS model accounts for the joint effects of the repeated measurements over the logistic curve (Eq. (1)). The implementation of the model was done by a maximum likelihood approach using the R package of fGWAS [6]. The Manhattan plot of the fGWAS is given in Figure 1. fGWAS analyses resulted in 22 associations with 12 false positives. Figure 2 shows how the genotypes of the most significant SNP (all_0.9137) affect the mean curve of the phenotype. The additive effects of all SNPs on the longitudinal phenotypes tend to increase with age (Figure 2).

3.2. Chicken data results

The mean heterozygosity (SD) for the SNPs and the chickens were predicted as 0.3398 (0.1436) and 0.3400 (0.0175), respectively. The estimated genomic heritabilities

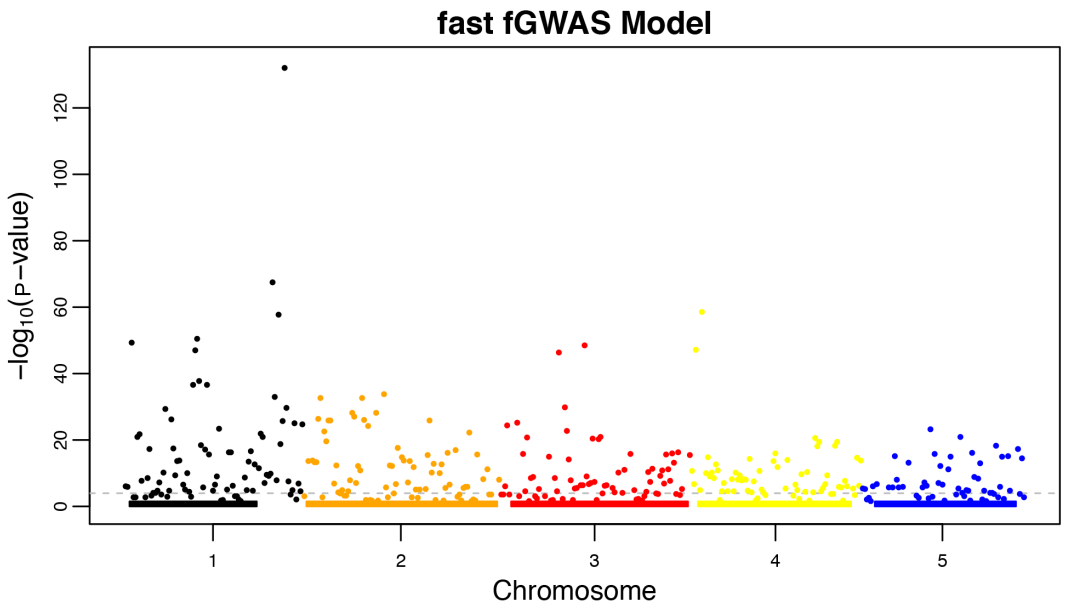

Figure 1. The Manhattan plots of –log 10 (P-values) using all SNPs obtained by the fGWAS model using QTLMAS dataset.

were 0.0176, 0.0068, and 1.20e-08 for the asymptote (A), inflection point (I), and scaling factor (S) parameters, respectively. The estimates of time specific heritabilities were 0.3638, 0.5205, 0.1058, 0.3769, 0.1078, and 0.1748 for the first EW and 28, 36, 56, 66, 72, and 80 weeks of age, respectively, using a pedigree-based restricted maximum likelihood approach [23].

The genomic inflation factors were predicted to be between 1 and 1.2160 by GRAMMAR and 1.03 and 1.33 by EGSCORE approaches. A genome-wide significant SNP affecting the asymptote was predicted on chromosome 2 and chromosome 6 using GRAMMAR (Table 2) (FDR corrected P-values of <0.0001) and EGSCORE (FDR corrected P-values of <0.0001) approaches (Table 3). SNPs affecting the inflection point were found in chromosome 2 using GRAMMAR (FDR corrected P-values of 0.0001) and EGSCORE (FDR corrected P-values of 0.0001) approaches. Two significant SNPs were detected by EGSCORE for scale parameter on chromosome 1 (FDR corrected P-values of 0.00013) and 11 (FDR corrected P-values of 0.0065). A genomic region on chromosome 2 in association with EW was also reported by Honkatukia

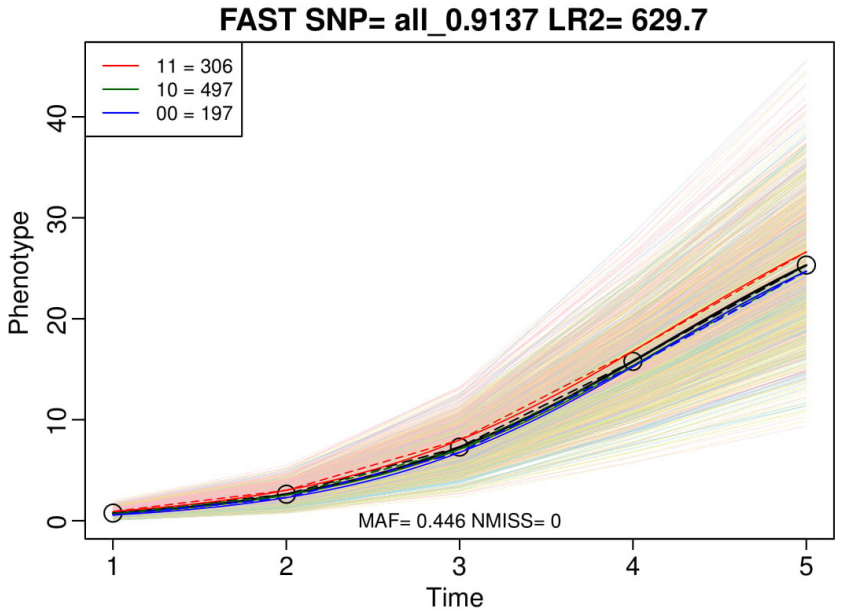

Figure 2. The mean logistic growth curve at the SNP all_0.9137 using 3 genotypes as 11, 10, and 00 with likelihood value of 629.7, minor allele frequency (MAF) of 0.446, and no missing genotypes (NMISS = 0).

Table 2. The GWA results obtained from the estimated asymptote parameters of logistics curve using GRAMMAR model for the chicken dataset.

Single nucleotide polymorphism	Chromosome	Location	No of individuals	Chi-square	P-val.	FDR
AX-75956307	2	104412633	1063	45.5958	1.45e-11	1.09e-06
AX-75956326	2	104421751	1063	45.5595	1.48e-11	1.09e-06
AX-75956389	2	104456371	1063	47.63451	5.14e-12	1.09e-06
AX-80755176	2	104444429	1063	45.5595	1.48e-11	1.09e-06
AX-80972053	6	32636946	1063	44.10385	3.11e-11	1.83e-06
AX-76950202	6	32596862	1063	40.77894	1.70e-10	3.58e-06
AX-76950231	6	32606739	1063	40.77894	1.70e-10	3.58e-06
AX-76950292	6	32624648	1063	40.77894	1.70e-10	3.58e-06
AX-76950312	6	32631539	1063	40.77894	1.70e-10	3.58e-06
AX-76950321	6	32634867	1063	40.77894	1.70e-10	3.58e-06

et al. [24] and Tuiskula-Haavisto et al. [25], similar to our results using the GRAMMAR and EGSCORE approaches.

The genetic architecture of longitudinal traits is related to the number and effects of associated SNPs conditional on the assumption for the shape of the mean curve of the repeated measurements of the phenotypes. We used fGWAS for a combination of the longitudinal EW measurements and genomic data to increase the power of the association analyses. fGWAS analyses detected 20 genome-wide significant SNPs (Table 4). Among them, 15 were located on chromosome 1. A Manhattan plot for fGWAS analyses is presented in Figure 3. Most of the significant SNPs are located within 168.71 and 168.86 Mb on GGA1.

We used Bayesian variable selection models with different priors to allow for SNPs with different effects in Bayes C(π) (Table 5) and BL (Table 6). One highly significant SNP (AX-76640358) was located on chromosome 4 on 2906524 bp by both Bayes C(π) and BL. Significant SNP effects on chromosome 4 were also reported but from different positions by Wolc et al. [26], Schreiweiss et al. [27], and Sasaki et al. [28]. The Bayes C(π) model using the asymptotic (a) revealed the strongest SNP association on chromosome 2 (AX-76096675). Several genomic regions on chromosome 2 were also located by Honkatukia et al. [24] on EW and Tuiskula-Haavisto et al. [25] on late EW (46–61 weeks).

4. Discussion

EWs have been considered among the important selection targets for both customers and producers, especially during certain laying periods [26,29]. In this study,

Table 3. The GWA results obtained from the estimated asymptote parameters of logistics curve using EGSCORE model for the chicken dataset.

Single nucleotide polymorphism	Chromosome	Location	No of individuals	Chi-square	P-val.	FDR
AX-75956073	2	104276426	1063	53.81052	2.21e-13	3.26e-08
AX-75956228	2	104369820	1063	53.81052	2.21e-13	3.26e-08
AX-75956230	2	104370166	1063	49.52012	1.96e-12	1.14e-07
AX-75956307	2	104412633	1063	48.94421	2.63e-12	1.14e-07
AX-75956326	2	104421751	1063	48.88544	2.71e-12	1.14e-07
AX-75956389	2	104456371	1063	49.11973	2.41e-12	1.14e-07
AX-80755176	2	104444429	1063	48.88544	2.71e-12	1.14e-07
AX-75956319	2	104419495	1063	48.4373	3.41e-12	1.22e-07
AX-76950027	6	32536761	1063	48.06352	4.13e-12	1.22e-07
AX-80853767	6	32530677	1063	48.06352	4.13e-12	1.22e-07

Table 4. Summary of top 10 identified loci for the chicken dataset using the fGWAS model.

SNP	CHR	Position	Minor allele frequency	P-val.
AX-80745839	1	168720658	0.425212	1.53e-08
AX-80826503	1	168719617	0.424271	1.79e-08
AX-75334447	1	168724175	0.423801	2.05e-08
AX-75334463	1	168729832	0.415804	5.84e-08
AX-75334594	1	168789601	0.417215	6.28e-08
AX-75334726	1	168844535	0.412512	1.51e-07
AX-75334678	1	168824266	0.412512	1.55e-07
AX-75334461	1	168728872	0.431797	1.65e-07
AX-75916442	18	9905356	0.069144	1.72e-07
AX-75334450	1	168724934	0.429915	1.72e-07

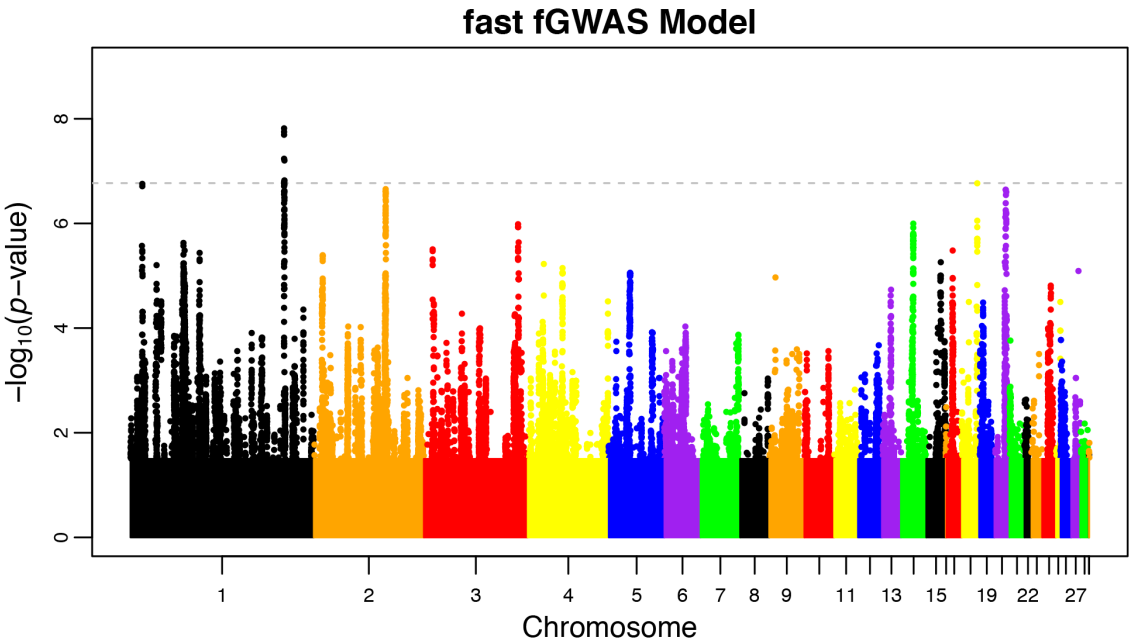

Figure 3. The Manhattan plots of –log 10 (P-values) using all SNPs obtained by the fGWAS model using chicken dataset.

Table 5. The GWA results obtained from the estimated asymptote parameters of logistics curve using Bayes C(π) for the chicken dataset.

SNP	Chromosome	Position	Effect
AX-76096675	2	45380463	64.90982
AX-76640358	4	2906524	50.52048
AX-76640730	4	2926392	46.38021
AX-76097075	2	45623951	44.50279
AX-76640820	4	2931026	43.00988
AX-76640646	4	2921304	40.50876
AX-76640535	4	2915510	39.79021
AX-76096597	2	45336622	36.45069
AX-77116057	8	6713905	35.72584
AX-77115981	8	6683020	35.06507

Table 6. The GWA results obtained from the estimated asymptote parameters of logistics curve using Bayesian lasso (BL) for the chicken dataset.

SNP	Chromosome	Position	Effect
AX-76640646	4	2921304	9.847588
AX-76365700	27	4149177	9.576412
AX-76365716	27	4153001	9.481536
AX-77116057	8	6713905	8.897123
AX-76640535	4	2915510	8.764108
AX-77115981	8	6683020	8.342496
AX-76640358	4	2906524	8.114362
AX-76365630	27	4129765	8.05678
AX-76365679	27	4144834	7.966066
AX-77115967	8	6676897	7.905176

we aimed to detect genomic variants that might affect longitudinal measurements of EWs using various genomic models. Two distinct assumptions were evaluated for the genetic architecture of the EW: (a) the use of the single SNP models to detect major genes, and (b) the use of whole genome regression models [11] to detect both small and major genes that might affect longitudinal EW measurements.

We propose that the analyses of the simulated datasets could be useful for comparing different statistical models. We investigated various genomic models using a simulated QTLMAS dataset. Our results show that Bayes C(π) (10/20) (Table 1) and fGWAS (12/22) (Figure 1) have greater accuracy than the other tested models. Our Bayes C(π) findings are in agreement with the results of Heuven and Janss [30] and Veerkamp et al. [31], who both used Bayes C type models and detected 8 and 9 QTLs from the QTLMAS dataset, respectively. Since the model does not include a polygenic component, we used a stringent false discovery rate (P < e-15) threshold for fGWAS.

Similarly, we observed falsely increased heritabilities when we excluded the polygenic component from model (1) as 0.5283, 0.8049, 0.8407, 0.5347, 0.7192, 0.4535, and 0.4238 for the first EW and 28, 36, 56, 66, 72, and 80 weeks of age, respectively. fGWAS has advantages over other single SNP models due to an improved power for association detection by achieving the combination of the longitudinal measurements over the mean trajectory [1].

The highest heritabilities were predicted for 28 (0.5205) and 36 (0.5027) weeks of age, which are consistent with the findings of Liu et al. [14]. Since Liu et al. [14] used genomic relationships to predict time specific heritabilities (we used pedigree-based variance components for obtaining heritability estimates), their heritability estimates decreased [32].

Using fGWAS, we mapped the associated region (Figure 3) previously shown to include a QTL affecting the EW on chromosome 1 [14]. Liu et al. [14] detected the genomic signal specifically for week 36 and our fGWAS data (when combining whole longitudinal observations) confirmed this result. However, Liu et al. [14] stated that the associated signals of chromosome 1 for week 36 also correlated with weeks 28, 56, and 66, but without reaching statistical significance.

The results obtained from this study show that fGWAS can be useful for manipulating dynamic EW over the entire laying period (based on a moderate to major effect). The fGWAS SNPs that were associated with EW were located on chromosome 1, close to the gene *DLEU7*, which regulates ovary weight in chickens [14,33]. The SNPs were detected based on the absolute effect sizes using Bayes C(π) (Table 5) and BL (Table 6) is likely useful for predicting the polygenic risk scores and/or genomic breeding values during genomic selection for longitudinal EW measurements.

Acknowledgement
This study was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK), Grant No: 118O108.

References

1. Das K, Li J, Wang Z, Tong C, Fu G et al. A dynamic model for genome-wide association studies. Human Genetics 2011; 129 (6): 629-639. doi: 10.1007/s00439-011-0960-6
2. Sikorska K, Lesaffre E, Groenen PJF, Rivadeneira F, Eilers PHC. Genome-wide analysis of largescale longitudinal outcomes using penalization GALLOP algorithm. Nature Scientific Reports 2018; 8: 6815. doi: 10.1038/s41598-018-24578-7
3. Wu RL, Lin M. Functional mapping how to study the genetic architecture of dynamic complex traits. Nature Reviews Genetics 2006; 7: 229-237. doi: 10.1038/nrg1804
4. Karacaören B. A Bayesian random walk approach for mapping dynamic quantitative trait. Journal of Applied Nonlinear Dynamics 2016; 5 (1): 105-115. doi: 10.5890/JAND.2016.03.008
5. Furlotte NA, Eskin E, Eyheramendy S. Genome-wide association mapping with longitudinal data. Genet Epidemiology 2012; 36: 463-471. doi: 10.1002/gepi.21640
6. Wang Z, Xu K, Zhang X, Wu X, Wang Z. Longitudinal SNP-set association analysis of quantitative phenotypes. Genetic Epidemiology 2017; 41 (1): 81-93. doi: 10.1002/gepi.22016
7. Wu Z, Hu Y, Melton PE. Longitudinal data analysis for genetic studies in the whole-genome sequencing era. Genetic Epidemiology 2014; 38 (S1): S74-S80. doi: 10.1002/gepi.21829
8. Rönnegård L, McFarlane SE, Husby A, Kawakami, Ellegren H et al. Increasing the power of genome wide association studies in natural populations using repeated measures –evaluation and implementation. Methods in Ecology Evolution 2016; 7 (7): 792-799. doi: 10.1111/2041-210X.12535
9. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA et al. Finding the missing heritability of complex diseases. Nature 2009; 461 (7265): 747-753. doi: 10.1038/nature08494
10. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome wide dense marker maps. Genetics 2001; 157: 1819-1829.
11. Campos G, Gianola D, Allison DB. Predicting genetic predisposition in human: the promise of whole-genome markers. Nature Reviews Genetics 2010; 11: 880-886. doi: 10.1038/nrg2898
12. Andersson L, Georges M. Domestic-animal genomics: deciphering the genetics of complex traits. Nature Reviews Genetics 2004; 5: 202-212. doi: 10.1038/nrg1294
13. Sikorska K, Montazeri NM, Uitterlinden A, Rivadeneira F, Eilers PH et al. GWAS with longitudinal phenotypes: performance of approximate procedures. European Journal of Human Genetics 2015; 23 (10): 1384. doi: 10.1038/ejhg.2015
14. Liu Z, Sun C, Yan Y, Li G, Wu G et al. Genome-wide association analysis of age-dependent egg weights in chickens. Frontiers in Genetics 2018; 9: 128. doi: 10.3389/fgene.2018.00128
15. Coster A, Bastiaansen JW, Calus MP, Maliepaard C, Bink MC. QTLMAS 2009: simulated dataset. BMC Proceedings 2010; 4 (Suppl. 1): S3. doi: 10.1186/1753-6561-4-S1-S3
16. West GB, Brown JH, Enquist BJ. A general model for ontogenetic growth. Nature 2001; 413: 628-631.
17. Ma CX, Casella G, Wu R. Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. Genetics 2002; 161 (4): 1751-1762.

18. Aulchenko YS, Ripke S, Isaacs A, Van Duijn CM. GenABEL: an R library for genome wide association analysis. Bioinformatics 2006; 23 (10): 1294-1296. doi: 10.1093/bioinformatics/btm108

19. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA et al. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 2006; 38 (8): 904-909. doi: 10.1038 / ng1847

20. Legarra A, Ricardi A, Filangi O. GS3: Genomic Selection, Gibbs Sampling, Gauss Seidel (and BayesCp). Paris, France: INRA; 2011.

21. Legarra A, Robert-Granié C, Manfredi E, Elsen JM. Performance of genomic selection in mice. Genetics 2008; 180: 611-618. doi: 10.1534/genetics.108.088575

22. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2018.

23. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T et al. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production; Montpellier, France; 2002.

24. Honkatukia M, Tuiskula-Haavisto M, de Koning DJ, Virta A, Maki-Tanila A et al. A region on chicken chromosome 2 affects both egg white thinning and egg weight. Genetics Selection Evolution 2005; 37: 563-577. doi: 10.1186/1297-9686-37-6-563

25. Tuiskula-Haavisto M, Honkatukia M, Vilkki J, de Koning DJ, Schulman NF et al. Mapping of quantitative trait loci affecting quality and production traits in egg layers. Poultry Science 2002; 81: 919-927. doi: 10.1093/ps/81.7.919

26. Wolc A, Arango J, Settar P, Fulton JE, O'Sullivan NP et al. Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. Animal Genetics 2012; 43: 87-96. doi: 10.1111/j.1365-2052.2012.02381.x

27. Schreiweis MA, Hester PY, Settar P, Moody DE. Identification of quantitative trait loci associated with egg quality, egg production, and body weight in an F2 resource population of chickens. Animal Genetics 2006; 37 (2): 106-112. doi: 10.1111 / j.1365-2052.2005.01394.x

28. Sasaki O, Odawara S, Takahashi H, Nirasawa K, Oyamada Y et al. Genetic mapping of quantitative trait loci affecting body weight, egg character and egg production in F2 intercross chickens. Animal Genetics 2004; 35 (3): 188-194. doi: 10.1111/j.1365-2052.2004.01133.x

29. Yi G, Liu W, Li J, Jiangxia Z, Qu L et al. Genetic analysis for dynamic changes of egg weight in 2 chicken lines. Poultry Science 2014; 93: 2963-2969. doi: 10.3382 / ps.2014-04178

30. Heuven HCM, Janss LLG. Bayesian multi-QTL mapping for growth curve parameters. BMC Proceedings 2010; 4 (Suppl. 1): S12-10. doi: 1186/1753-6561-4-s1-s12

31. Veerkamp RF, Verbyla K, Mulder HA, Calus MPL. Simultaneous QTL detection and genomic breeding value estimation using high density SNP chips. BMC Proceedings 2010; 4 (Suppl. 1): S9-10. doi: 1186/1753-6561-4-s1-s9

32. Veerkamp RF, Mulder HA, Thompson R, Calus MPL. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. Journal of Dairy Science 2011; 94 (8): 4189-4197. doi: 10,3168 / jds.2011-4223

33. Sun C, Lu J, Yi G, Yuan J, Duan Z et al. Promising loci and genes for yolk and ovary weight in chickens revealed by a genome-wide association study. PLoS One 2015; 10: e0137145. doi: 10.1371/journal.pone.0137145