

A genomic investigation on hybrid sterility in house mouse using a two stage model

Burak KARACAÖREN

University of Akdeniz, Faculty of Agriculture, Department of Animal Science, Section of Biometry and Genetics, Antalya, Turkey.

Summary: Genome wide association studies (GWASs) commonly used to search for genetic variants associated with quantitative traits. Pleiotropic effect of genes may cause the observed correlations among different phenotypes. This study proposed a two stage multilocus model for pleiotropic GWAS using a Bayesian mixture model to take into account of both small and major gene effects. The objectives of this study were to investigate if the two-stage model was useful for detecting pleiotropic genes using a simulated data set and to investigate existence of pleiotropic genes for testis weight and testis gene expression levels in house mouse. The analyses included relative testis weight and testis gene expression traits. The results showed that two stage model had higher power to detect the pleiotropic QTL than the single marker model. It was also noted the possible economical impact of sampling informative individuals for the GWAS analyses by observing genomic trends in the simulated dataset. Two stage model detected 50 and 53 major SNP effects using first and the second principal components. Additive genetic variation explained by chromosome X was found to be 4% for the testis weight.

Keywords: Bayesian mixture model, genome wide association analyses, pleiotropy.

İki aşamalı bir model kullanarak ev faresinde hibrit sterilite üzerine genomik bir araştırma

Özet: Genom tabanlı ilişki analizleri (GTİÇ) kantitatif verimler ile ilgili genleri bulmak için sıklıkla kullanılmaktadır. Farklı fenotipler arasındaki korelasyon pleiotropik genlerce açıklanabilir. Bu çalışmada, pleiotropik GTİÇ'da hem mendelci hem de major etkili genlerin iki aşamalı çoklu lokus modeli sayesinde birlikte kullanılmaları önerilmektedir. Bu çalışmanın ana amacı: bir benzeşim veri seti kullanarak: çoklu lokus yaklaşımın pleiotropik genlerin bulunmasında faydalı olup olmadığının incelenmesi ve bir fare verisinde bu modeli kullanarak testis ağırlığı ve gen ifadelerinden sorumlu küçük ve major etkili pleiotropik genleri tespit etmektir. Göreceli testis ağırlıkları ve testis gen ifadesi karakterleri incelenmiştir. Sonuçlar iki aşamalı modelin pleiotropik QTL'leri tesbit etmede, tekil işaretleyici modelden daha yüksek güce sahip olduğunu ortaya koymuştur. Ek olarak benzeşim veri setlerinin genomik eğilimlerinin incelenmesi: GTİÇ'da bilgi veren bireylerin seçilmesinin daha ekonomik olduğunu ortaya koymuştur. Birinci ve ikinci temel bileşenli iki aşamalı model ile 50 ve 53 adet major SNP tesbit edilmiştir.. X kromozomunun, testis ağırlığındaki eklemeli genetik çeşitliliğin %4'ünü açıkladığı bulunmuştur.

Anahtar sözcükler: Bayeşçi karışım modeli, genom tabanlı ilişki analizi, pleiotropi.

Introduction

Gene flow is one of the important sources of genetic variation by hybridization of two distinct populations. Genetic biodiversity increase at hybridization area of locally separated populations (21). This hybrid might have hybrid sterility phenotypes due to divergent alleles at various loci (25). Storchova et al. (20) investigated the role of X chromosome in association with hybrid sterility phenotype.

Genome wide association studies (GWASs) commonly used to search for genetic loci associated with quantitative phenotypes. A single gene might affects multiple phenotypes: that leads to construct of networks and pathways of genes (7,8). Pleiotropic effect of genes may cause the observed correlations among different

phenotypes (23). Hartley et al (8) reported more than 8500 associations obtained from 350 human complex traits: indicates that pleiotropy exist (24). Shriver (18) reviewed the statistical models to detect pleiotropic effects. Currently multivariate approaches are commonly used to combine multiple phenotypes for pleiotropic GWAS analyses. Gao et al. (4) demonstrated that GWAS using residuals of linear combination of multiple phenotypes lead to higher statistical power and lower false discoveries over single trait analyses. However most of the pleiotropic models are employing single marker regression approaches using outcome of multivariate models. Assumptions for different gene effects lead to different models (hence results) in GWAS. Yang et al. (26) used a polygenic GWAS approach for human height and

explained higher proportion of missing heritability (explanatory variance was found to be 50 %) compared with the common practice of single marker regression GWAS (explanatory variance was found to be 5 %). Hill (9) suggested to incorporate small and pleiotropic effects into genomic studies. Gratten et al (7) also stressed the importance of both polygenic and mendelian pleiotropic effects for the proper interpretation of GWAS. Based on this suggestions and discussions this study proposed a two-stage model for pleiotropic GWAS using a Bayesian mixture model (16) to take into account of both small (polygenic) and major gene effects.

The objectives of this study were to 1) investigate if the two-stage model was useful for detecting pleiotropic genes using a simulated data set (22) and 2) to investigate existence of associated loci and pleiotropic genes for hybrid sterility (for testis weight and testis gene expression levels) for house mouse (21) using the two stage model.

Materials and Methods

The present work describes the two-stage model for detecting pleiotropic genes using both simulated QTLMAS data and real mouse datasets.

QTLMAS data set: The pedigree included four generations with 4100 individuals for the simulated quantitative traits (22). The genome consisted of 10000 Single Nucleotide Polymorphisms (SNPs) distributed over 5 chromosomes. The number of population founders were 1020 (1000 females and 20 males). Each male mated with 51 females to create families. Generations were forced to be nearly discrete hence overlapping. Phenotypes of 3080 individuals were available. Three correlated quantitative phenotypes were simulated to mimic milk yield (heritability =0.35), fat yield (heritability=0.35) and fat content (heritability=0.50) in the sheep. The detail of the simulated dataset could be found in Usai et al (22).

Mice Data Set: The mice data consists of 179 males from the house mouse hybrid zone (21). The analyses included relative testis weight and testis gene expression traits. Before statistical analyses, SNPs were removed from the data if the call rate are < 90%, minor allele frequencies are < 1 %. Finally, a total of 152342 SNPs were collected for the genomic analyses. The detail of the mouse dataset could be found in Turner and Harr (21).

A two Stage Pleiotropic GWAS model: This study proposed a two-stage model of pleiotropic GWAS as a means of increasing power of detecting both small and major gene effects using multivariate dimension reduction (3) and genomic prediction models (15). Following we give a general formulation of the two- stage model.

Factor and principal component analyses could be used to discover loadings that combines underlying

relationships hence pleiotropy among different phenotypes in GWAS. Stage 1 can be based on linear and orthogonal combinations of the residuals of the correlated phenotypes (y_1, y_2, \dots, y_p);

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ &\vdots \\ y_p &= a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pp}x_p \end{aligned}$$

with the coefficients (a) being chosen to convey highest explanatory variances compared with the original explanatory variables, x_1, x_2, \dots, x_p , (3).

The second stage employs a Bayesian mixture model (BayesR) (16) for predicting pleiotropic SNP effects based on outcome of the Stage 1. BayesR assumed a hierarchical models of mixtures using different SNP effects;

$$p(\beta_j | p, \sigma_g^2) = \sum_{i=1}^{k=4} p_i f(x | \theta_i)$$

where β is the SNP effects, p is the mixture proportions (assumed to be 0.00001, 0.0001, 0.001, 0.01), σ_g^2 is the genetic variance, $f(x | \theta_i)$ is normally distributed mixture densities with θ parameters vectors and observations, x , . The present study sampled 50000 markov chains and discarded first 5000 as burn in period and recorded every 10th sample for thinning the chain. Uninformative priors were also used to obtain desired posteriors.

A Single SNP Model: Discordant sib pair experimental design was also employed as described in Karacaören (10) using Yu's linear mixed model (27) as implemented in rrBLUP package (2,17). Pearson homogeneity statistic could be estimated from 2xm table via following formula;

$$T^1 = \sum_{j=1}^m \frac{(n_{1j} - n_{2j})^2}{n_{1j} + n_{2j}}$$

where n_{ij} stands for counted alleles among cases and controls, $i = 1, 2$ for sib 1 and sib 2, respectively and $j = 1 \dots m$ (number of alleles). Due to dependency of test statistics within siblings; permutation tests could be used for deciding of statistical thresholds. This study permuted the case and control states of the siblings with probability of 1/2 to detect threshold of significance level of the test statistics.

The present study used genomic kinship information in linear mixed model to take into account of the pedigree structure as was implemented in rrBLUP: (2) a linear mixed model (27) approach in R software (17)

The linear mixed model (single marker regression approach) was used as

$$y = Xb + Za + e \quad (1)$$

where y contains the observations, b is the sex effects, a is the additive genetic effect, matrices X and Z are incidence matrices, and e is a vector containing residuals.

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

For the random effects, it is assumed that \mathbf{A} is the coefficient of coancestry obtained from genotype of animals; \mathbf{I} is an identity matrix, σ_a^2 is the additive genetic variance and σ_e^2 is the residual variance.

Results

QTLMAS dataset: 301 SNP's were excluded due to minor allele frequency <1%, leaving 9699 SNP's in the analyses. Mean heterozygosity for the SNPs and the individuals were estimated as 0.35 with a standard error of 0.01. Normality was confirmed by Kolmogorow Smirnow test, $p > 0.150$ for each trait. Based on model (1) estimated genomic heritabilities were 0.31, 0.30, 0.47 for quantitative traits, respectively.

Different genetic architectures were assumed for the QTLMAS phenotypes. BayesR assumed that SNPs effects were obtained by different mixture proportions from the normal distribution. Since some phenotypes could be controlled by both polygenic and major genes such an approximation might be useful. The Markov chain algorithm was run 4 times and the trace plot of number of significant SNPs was investigated (Table 1). Visual inspection of the trace plots show convergence (results not

shown) of Markov chains. The posterior mean number of major SNPs were 51, 73 and 200 for trait 1, trait 2 and trait 3 respectively.

Single marker regression model was used (2, 27) to detect putative genomic associations using different number of discordant sib pairs (DSP) (Table 1). DSP design leads to homogenized samples by controlling for population stratification. In DSP design, sib pairs from each family could be used as cases and controls. Boehnke and Langefeld (1) used one case and one control from each family. In this study different number of DSP was sampled to association mapping using QTLMAS simulated dataset similar to Karacaören et al. (11) and Karacaören(12). Table 1 showed that number of detected true SNPs and QTLs increased with larger discordant sib pair families.

Table 2 (a, b ,c, d) presents genotype and allele counts for top two markers (1685 and 3585) by counting all alleles (scheme 1) or discordant alleles (scheme 2). There was concordant results for both markers using both allele counting schemes and GWA analyses. The results showed that both GWA using all individuals and samples of discordant sib pairs gave similar results. Contrasting alleles that are discordant between sib pairs (scheme 2) also increased the association test statistics compared with test statistics obtained by all alleles (scheme 1) (Table 3). The p values were obtained using Monte Carlo simulations (Figure 1) based on 100.000 permutations of the data. Permutated p -values show agreement with GWA results.

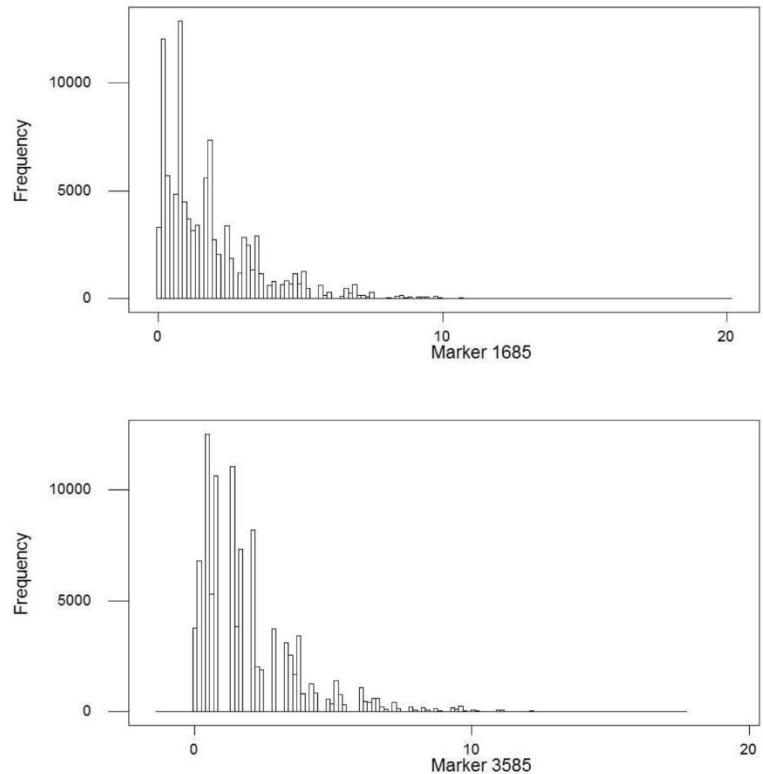


Figure 1. Permutation results for markers 1685 and 3585.

Şekil 1. 1685 ve 3585. işaretleyiciler için permutasyon sonuçları

Table 1. Number of true / number of total significant SNPs (number of true QTLs) from genome wide association analyses (GWA) with all individuals (All) and different sampling schemes for discordant sib pairs (DSP) ($n=2, 5, 10, 15, 20$).

Tablo 1. Gerçek/ toplam önemli SNP'lerin bütün bireyler (All) ve farklı örnekleme desenleri ile ayrık kardeş deneme deseni için (DSP) ($n=2, 5, 10, 15, 20$) genom tabanlı ilişki analizi (GWA) sonuçları.

	Trait 1	Trait 2	Trait 3
($n=2$)	7/8 (4)	3/7 (3)	14/18 (7)
($n=5$)	22/23 (11)	12/26 (6)	22/23 (11)
($n=10$)	24/24 (12)	16/39 (8)	16/25 (8)
($n=15$)	30/34 (15)	18/34 (9)	23/36 (11)
($n=20$)	32/32 (16)	16/34 (8)	22/37 (11)
All	22/32 (11)	16/35 (8)	22/40 (11)
BAYESR	32/51 (16)	28/73 (14)	56/200 (28)

Table 2a. Genotype counts for markers 1685

Tablo 2a. 1685 işaretleyicisi için genotip sayımları.

DSP GENOTYPE COUNT=1685			
	AFFECTED-SIB GENOTYPE		
UNAFFECTED-SIB GENOTYPE	AA	AB	BB
AA	6	5	1
AB	6	18	5
BB	6	10	3

Table 2b. Genotype counts for markers 3585

Tablo 2b. 3585 işaretleyicisi için genotip sayımları.

DSP GENOTYPE COUNT=3585			
	AFFECTED-SIB GENOTYPE		
UNAFFECTED-SIB GENOTYPE	AA	AB	BB
AA	0	2	0
AB	6	14	5
BB	4	15	14

Table 2c. Allele counts for markers 1685

Tablo 2c. 1685 işaretleyicisi için alel sayımları.

DSP ALLELE COUNT=1685		
	A	B
All Alleles (scheme 1)		
Affected Sibs	53	67
Unaffected Sibs	69	51
Discordant Alleles (scheme 2)		
Affected Sibs	18	33
Unaffected Sibs	33	22

Table 2d. Allele counts for markers 3585

Tablo 2d. 3585 işaretleyicisi için alel sayımları.

DSP ALLELE COUNT=3585		
	A	B
All Alleles (scheme 1)		
Affected Sibs	29	91
Unaffected Sibs	51	69
Discordant Alleles (scheme 2)		
Affected Sibs	13	34
Unaffected Sibs	31	22

Table 3. Test statistics, T^1 , for all alleles, DSP alleles and whole genome wide association analyses for different markers (Pearson correlation coefficient values and p values)

Tablo 3. Bütün alleler DSP alleleri ve bütün genomik ilişki analizindeki farklı işaretleyiciler için test istatistikleri, T^1 (Pearson korelasyon katsayısı değerleri ve p değerleri).

	All Alleles	DSP allele count	$-\log(P)$
Marker 1685	4.27 (0.10094)	6.62 (0.02886)	17.17
Marker 3585	12.86 (0.00036)	15.71 (0.00003)	14.47

Table 4. Estimates of genomic heritabilities using first (1) and second (2) scores obtained from different rotation techniques.

Tablo 4. Farklı döndürüm (rotasyon) tekniklerine ait birinci (1) ve ikinci (2) puanlardan elde edilmiş genomik kalıtım derecesi tahminleri.

Model	Estimates of heritability
Principal Component 1	0.33
Principal Component 2	0.45
Harris-Kaiser 1	0.31
Harris-Kaiser 2	0.36
Promax 1	0.31
Promax 2	0.37
Varimax 1	0.32
Varimax 2	0.37

Linear combinations of multivariate phenotypes were used (loadings) for detecting pleiotropy. Loadings were used as a response variable with the models to detect pleiotropic genes. Principal component analyses can be used to discover underlying pleiotropic patterns of genomic datasets. The method could also be used for reduction of dimension of datasets as well. Principal component rotation techniques were used to obtain linear combinations of multivariate phenotypes for detecting pleiotropy. Loadings were used to estimate heritabilities and to detect putative associations (Table 1 and Table 4).

Genomic heritabilities were estimated based on first two loadings from principal components and factor analyses (Table 4) with different rotations (16). The highest heritability was obtained by second principal components (0.45). However first principal component had the highest explanatory variation (0.620 and 0.376 respectively). Hence both principal component scores were used to detect pleiotropy in subsequent analyses.

Genome scans of the principal components were generally successful in finding significant true positive genomic signals for pleiotropy. In single marker regression analyses for each of the two top principal components: 28 (of 33) and 20 (of 33) SNPs were detected using the linear mixed model. In two stage analyses of each of the two top principal components: 20 (of 48) and

30 (45) SNPs were detected. A total of 36 and 48 pleiotropic QTLs were detected by single marker and the two stage approaches. Therefore, the two stage model had higher power to detect the pleiotropic QTL than the single marker model.

Mouse data set: A hierarchical Bayesian mixture model was used (16) for detecting major and minor SNP effects for relative testis weight and genome wide testis gene expression patterns. Since most of the quantitative traits influenced by number of genes with various effect sizes: such an assumption might lead to useful findings in practice (6). Visual inspection of the trace plots show convergence (results not shown) of Markov chains. The posterior mean number of SNPs for the testis weight were 1123.

Table 5 shows that there are SNPs with major SNP effects. Additive genetic variation explained by chromosome X was found to be 4% for the testis weight. Over all 1123 SNPs explained 99% of the total genetic variance.

Table 5. Predicted SNPs effects for testis weight by a Bayesian mixture model.

Tablo 5. Testis ağırlığı için Bayeşçi karışım modeli ile elde edilmiş SNP etkileri

SNP	Chromosome	Position	Effect
JAX00486833	2	33426476	0.056959
JAX00034959	12	37595107	0.036166
JAX00391946	14	124643672	0.025451
JAX00596027	5	127055315	0.024441
JAX00397676	15	34048660	0.021275
JAX00659084	7	151799616	0.019798
JAX00664173	8	29420605	0.019018
JAX00639799	7	62418975	0.019005
JAX00273087	1	160478940	0.018338
JAX00486833	2	33426476	0.056959

To identify fertility related pleiotropic genomic regions in 179 male mice: relative testis weight (testis weight/ body weight) and genome wide testis gene

expression patterns were analysed using the two-stage model (Table 6). The first principal component account for highest explanatory proportion (14 %) of the data. Each succeeding principal components predicted very small sizes of SNPs effects. Two-stage model detected 50 and 53 major SNP effects using first and the second principal components.

Discussion and Conclusion

QTLMAS data set: Table 1 showed that number of detected true SNPs and QTLs increased with larger discordant sib pair families. This was particularly true for the trait 1. Although Marker 1685 had highest $-\log P$ value (Table 3), it was not found to be statistically significant by counting Scheme of 1 ($p = 0.1009$). The main reason was uncorrected inbreeding structure in the genotypic data set. However, when counting Scheme of 2 was used to incorporate the inbreeding structure, the genomic signal was detected by at the level of 100.000 permutations. There was a good agreement for marker 3585 using both allele counting schemes and whole genome wide association analyses (Table 3). The results showed that both GWA using all individuals and samples of discordant sib pairs gave similar results. Contrasting alleles that are discordant between sib pairs (scheme 2) also increased the association test statistics compared with test statistics obtained by all alleles (scheme 1) (Table 3). Hence evidence for association was much higher when using discordant alleles instead of using all alleles. The p values were evaluated using Monte Carlo simulations (Table 3) based on 100.000 permutations of the data. Permuted p-values show agreement with full genome wide association results. Stronger association were observed for marker 3585 compared with marker 1685. The possible economical impact of sampling informative individuals was also noted for the GWAS analyses by observing genomic trends in the Table 1. For example, usage of discordant sib pairs of 20 (detected 16 QTLs) had higher accuracy instead of using the whole individuals (detected 11 QTLs).

Table 6. Summary of top ten identified pleiotropic loci of the mouse data set.

Tablo 6. Fare veri setinden elde edilmiş ilk on pleiotropik lokusa ait özet.

PRINCIPAL COMPONENT	SNP ID	CHR	POSITION	EFFECT SIZE
1	JAX00498433	2	103801050	167.168
1	JAX00266356	1	130401319	95.2271
1	JAX00598774	5	137616662	74.6396
1	JAX00706991	9	117110944	72.2723
1	JAX00611301	6	58617650	64.7165
1	JAX00050167	14	13585763	63.1884
1	JAX00488905	2	49260698	53.7426
1	JAX00263226	1	115832554	53.1671
1	JAX00309656	11	45140995	33.1346
1	JAX00661579	8	14496838	26.6909

The two stage analyses were able to identify additional new positions missed by the single marker regression model. This was particularly true for the phenotype 2 (14 true QTLs) and phenotype 3 (28 true QTLs).

For pleiotropic analyses: simulated GWAS data of 9699 SNPs were employed with 3080 individuals for 3 quantitative phenotypes: the heritabilities were equal to 0.35, 0.35, and 0.50 respectively. The outcome of principal component analyses of the 3 quantitative phenotypes were used for detecting pleiotropic effects by the Bayesian hierarchical model (BayesR). The assumed SNPs effects were obtained with different effect sizes: the first containing 56 SNPs with major effects, the second containing 128 SNPs with smaller effect, and a large group of 1410 SNPs representing a polygenic component for the first principal component. The second principal component contained 64 major effects, 66 moderate effects and 545 polygenic effects. Since the QTLMAS data set simulated with high level of inbreeding structure: moderate and polygenic effects were essential for taking into account of the associated pedigree (and population) structure. The genomic signals from the major effects of the BayesR were compared with the organizers true results (22) by the threshold of 1Mb. BayesR was able to detect 20 and 30 true SNPs using first and second principal components respectively. A total of 48 (out of 50) QTLs were detected on both principal components. In general, the two-stage model showed the highest detection power, particularly when it was used with the second principal component. However, the two stage model also produced larger number of false positives: success rates were found to be 0.42 and 0.67 (for principal components 1 and 2, respectively): compared with the results of single marker regression approach as 0.85 and 0.61 (for principal components 1 and 2, respectively). However, the two-stage model was able to detect higher number of true QTLs (48 true QTLs) compared with the single marker regression approach (36 true QTLs). Similar to our findings, importance of genetic architecture was also stressed by Leong et al. (13), Liu et al. (14) and Usai et al. (22) by comparing the different genetic models and assumptions for detecting pleiotropic genes. However, Gianola et al. (5) pointed out potential problems for detecting pleiotropy using molecular markers due to unaccounted linkage disequilibrium structure in the GWAS. The proposed two-stage model may have limitations in terms of false positive signals, however it has an advantage over most other single marker regression models since it does take into account of linkage disequilibrium by modeling whole SNPs simultaneously.

Mouse data set: Turner and Harr (21) assumed major genes for the testis weight and detected genomic signals from 9 autosomes and X chromosome by using permissive

false discovery rate (stringent threshold detected four SNPs). Since Turner and Harr (21) used threshold for p values to take into account of multiple hypothesis testing such results expectable. The proposed model that affects multiple SNPs simultaneously can resolve the multiple hypothesis testing problem in pleiotropic analyses of GWAS. BayesR model provides flexibility in terms of effect sizes for the SNPs. The analyses of this study showed that most of the SNPs (>0.98) had small effects for the testis weight, demonstrating importance of the polygenic component. By carrying out a pleiotropic GWAS in a male house mouse, up to 103 SNPs that are associated in top two principal components variation with joint effects of testis weight and testis gene expression levels were identified. Top two principal components explain 21 % of the total variation.

A new two stage method to detect pleiotropic SNPs using a Bayesian mixture model (16) was presented to increase accuracy in this study. By definition, GWAS uses hypothesis free explanatory tools. In that regard two-stage model of the present study might be useful to investigate pleiotropic GWAS referring different genetical architectures for the phenotypes. Employing the Bayesian mixture model (BayesR) at the second stage corrects for both the linkage disequilibrium and the polygenic effects (16). Compared with the single marker regression pleiotropic models: the two stage model was also prone to well known multiple hypothesis testing problems in GWAS. Although the simulated data showed that two-stage model of the present study subject to false positive results, it still might be informative (and hypothesis producer) to use it for scanning the genome to detect major and small pleiotropic effects.

Acknowledgements

The study was supported by Akdeniz University Scientific Research Projects Unit, project number: FBA-2015-1117.

References

1. **Boehnke M, Langefeld CD** (1998): *Genetic association mapping based on discordant sib pairs: the discordant-alleles test*. *Am J Hum Genet*, **62**, 950-961.
2. **Endelman JB** (2011): *Ridge regression and other kernels for genomic selection with R package rrBLUP*. *The Plant Genome*, **4**, 250-255.
3. **Everitt BS, Landau S, Leese M** (2001): *Cluster Analysis*. National Academy Press, Washington, DC
4. **Gao H, Zhang T, Wu Y, et al** (2014): *Multiple-trait genome-wide association study based on principal component analysis for residual covariance matrix*. *Heredity*, **113**, 526-532.
5. **Gianola D, de los Campos G, Toro, MA, et al** (2015). *Do molecular markers inform about pleiotropy?* *Genetics*, **201**, 23-29.

6. **Goddard, ME, Kemper, KE, Macleod, IM, et al** (2016): *Genetics of complex traits: prediction of phenotype, identification of causal polymorphisms and genetic architecture*. In Proc. R. Soc. B, **283**, 20160569.
7. **Gratten J, Visscher PM** (2016): *Genetic pleiotropy in complex traits and diseases: implications for genomic medicine*. Genom Med, **8**, 78.
8. **Hartley SW, Monti S, Liu CT, et al** (2012): *Bayesian methods for multivariate modeling of pleiotropic SNP associations and genetic risk prediction*. Front genet, **3**, 176.
9. **Hill WG** (2010): *Understanding and using quantitative genetic variation*. Philosophical Transactions of the Royal Society of London B: Biological Sciences, **365**, 73-85.
10. **Karacaören B** (2016): *Investigations on Genetic Architecture of Hairy Loci in Dairy Cattle by Using Single and Whole Genome Regression Approaches*. Asian-Australas j anim sci, **29**, 938.
11. **Karacaören B, de Koning DJ, Velander I, et al** (2010): *Alternative association analyses on boar taint using discordant sib pairs experimental design*. In 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. 743 p.
12. **Karacaören B.** (2012): *Some observations for discordant sib pair design using QTL-MAS 2010 dataset*. Kafkas Vet Fak Derg, **18**, 857-860.
13. **Leong W, Sun PY, Edmands S.** (2017): *Latitudinal clines in temperature and salinity tolerance in tidepool copepods*. J. Hered, **109**, 71-77.
14. **Liu J, Wan X, Ma S, Yang C.** (2016): *EPS: an empirical Bayes approach to integrating pleiotropy and tissue-specific information for prioritizing risk genes*. Bioinformatics, **32**, 1856-1864.
15. **Meuwissen THE, Hayes BJ, Goddard ME** (2001): *Prediction of total genetic value using genome wide dense marker maps*. Genetics **157**, 1819-1829.
16. **Moser G, Lee HS, Hayes BJ, et al** (2015): *Simultaneous discovery, estimation and prediction analysis of complex traits using a bayesian mixture model*. PLoS. Genet. **11**, e1004969.
17. **R Development Core Team** (2013): *R: A language and environmental for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
18. **Shriner D** (2012): *Moving toward system genetics through multiple trait analysis in genome-wide association studies*. Front Genet, **3**, 1.
19. **Solovieff N, Cotsapas C, Lee PH, et al** (2013): *Pleiotropy in complex traits: challenges and strategies*. Nat Rev Genet, **14**, 483-495.
20. **Storchova R, Gregorova S, Buckiova D, et al** (2004): *Genetic analysis of X-linked hybrid sterility in the house mouse*. Mamm. Genome, **15**, 515-524.
21. **Turner LM, Harr B.** (2014): *Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions*. Elife, **3**, e02504.
22. **Usai MG, GaspaG, Macciotta NP, et al** (2014): *XVI th QTLMAS: simulated dataset and comparative analysis of submitted results for QTL mapping and genomic evaluation*. In BMC proceedings (Vol. 8, No. 5, p. S1). BioMed Central.
23. **Visscher PM, Yang J** (2016): *A plethora of pleiotropy across complex traits*. Nat Genet, **48**, 707-708.
24. **Wang L, Oehlers SH, Espenschied ST, et al** (2015): *CPAG: software for leveraging pleiotropy in GWAS to reveal similarity between human traits links plasma fatty acids and intestinal inflammation*. Genome biol, **16**, 190.
25. **White, MA, Stubbings M, Dumont BL, et al** (2012): *Genetics and evolution of hybrid male sterility in house mice*. Genetics, **191**, 917-934.
26. **Yang J, Benyamin B, McEvoy BP, et al** (2010): *Common SNPs explain a large proportion of the heritability for human height*. Nat Genet, **42**, 565-569.
27. **Yu J, Pressoir G, Briggs WH, et al** (2006): *A unified mixed-model method for association mapping that accounts for multiple levels of relatedness*. Nat genet, **38**, 203-208.

Geliş tarihi: 22.05.2017 / Kabul tarihi: 20.10.2017

Address for correspondence:

Doç. Dr. Burak KARACAÖREN
Akdeniz Üniversitesi, Ziraat Fakültesi
Zootečni Bölümü, Biyometri ve Genetik Kürsüsü
07059, Antalya TÜRKİYE
burakkaracaoren@akdeniz.edu.tr