



Multiple Genes Related to Muscle Identified through a Joint Analysis of a Two-stage Genome-wide Association Study for Racing Performance of 1,156 Thoroughbreds

Dong-Hyun Shin¹, Jin Woo Lee², Jong-Eun Park¹, Ik-Young Choi³, Hee-Seok Oh⁴,
Hyeon Jeong Kim⁵, and Heebal Kim^{1,5,*}

¹ Department of Agricultural Biotechnology, Animal Biotechnology Major, and
Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Korea

- Supplementary Data -

Supplementary Figures

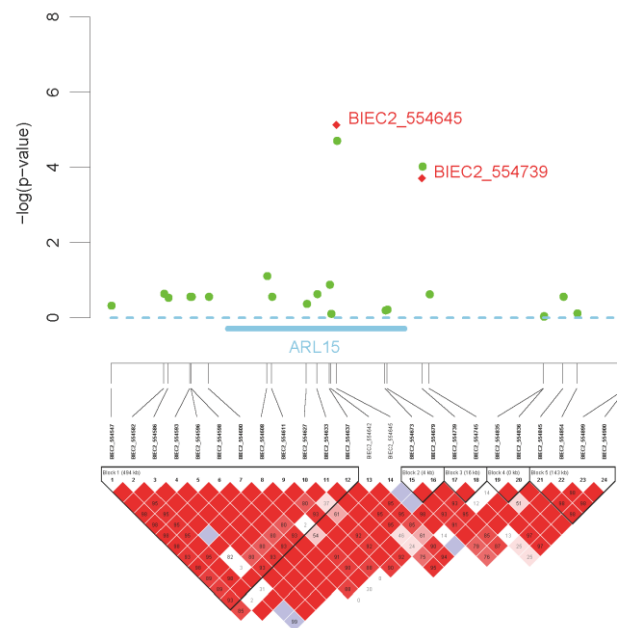
A. Supplementary Figure S1.	p.2
B. Supplementary Figure S2.	p.3
C. Supplementary Figure S3.	p.4
D. Supplementary Figure S4.	p.5
E. Supplementary Figure S5.	p.6
F. Supplementary Figure S6.	p.7
G. Supplementary Figure S7.	p.8
H. Supplementary Figure S8.	p.9
I. Supplementary Figure S9.	p.10
J. Supplementary Figure S10.	p.11
K. Supplementary Figure S11.	p.12

Supplementary Tables

L. Supplementary Table S1.	p.13
---------------------------------	------

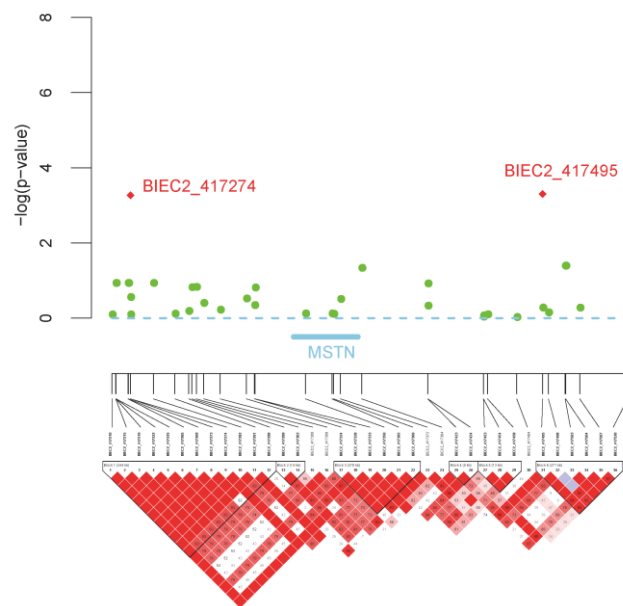
Supplementary Figure S1

Location of the association signal and pairwise linkage disequilibrium (LD) surrounding two on chromosome 21. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, $n = 240$); significant SNPs in combined dataset of stage 2 (red diamond, $n = 1,156$). This region has *ARL15* genes related to two significant SNPs.



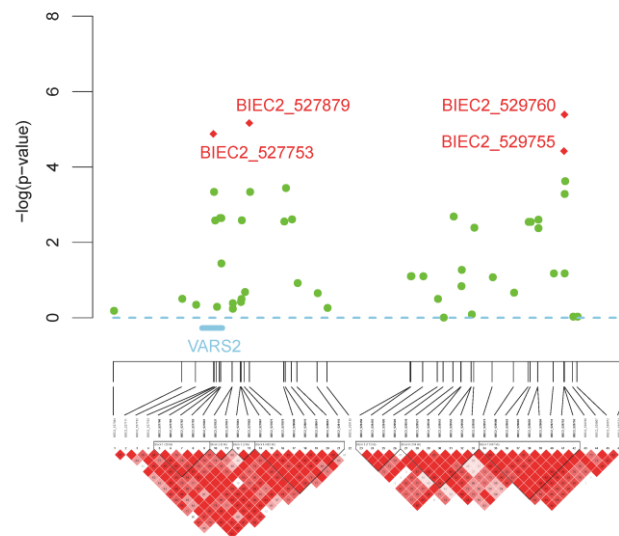
Supplementary Figure S2

Location of the association signal and pairwise linkage disequilibrium (LD) surrounding four on chromosome 21. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, $n = 240$); significant SNPs in combined dataset of stage 2 (red diamond, $n = 1,156$). This region has *CCT5*, *TAS2R1* genes related to each two, one significant SNPs.



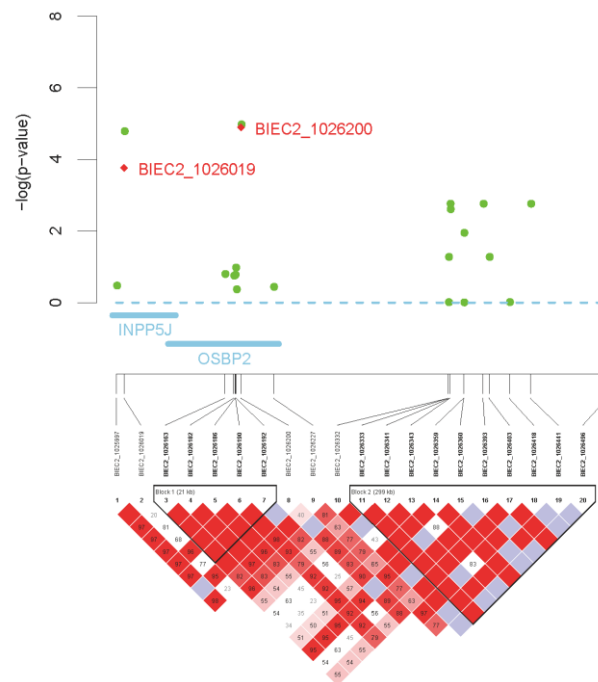
Supplementary Figure S3

Location of the association signal and pairwise linkage disequilibrium (LD) surrounding four on chromosome 20. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, $n = 240$); significant SNPs in combined dataset of stage 2 (red diamond, $n = 1,156$). This region has *VAR2* genes related to one significant SNPs.



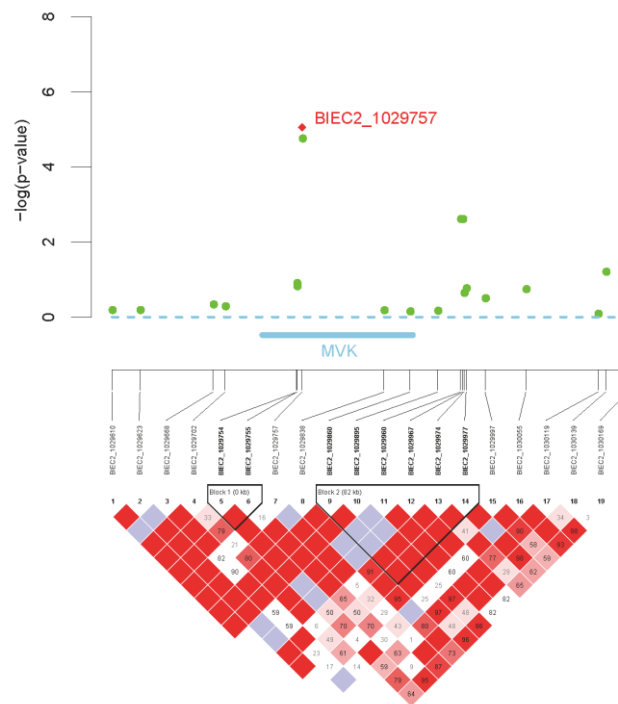
Supplementary Figure S4

Location of the association signal and pairwise linkage disequilibrium (LD) surrounding four on chromosome 8. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, $n = 240$); significant SNPs in combined dataset of stage 2 (red diamond, $n = 1,156$). This region has *INPP5J*, *OSBP2* genes related to each one significant SNP.



Supplementary Figure S5

Location of the association signal and pairwise linkage disequilibrium (LD) surrounding four on chromosome 8. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, $n = 240$); significant SNPs in combined dataset of stage 2 (red diamond, $n = 1,156$). This region has *MVK* genes related to one significant SNP.

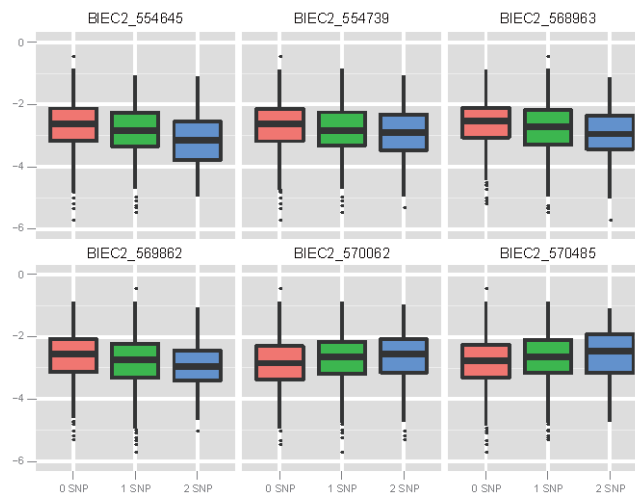


Location of the association signal and pairwise linkage disequilibrium LD) surrounding four on chromosome 30. This LD pattern is depicted using stage 1 data. Association signals are shown for all SNPs genotyped in stage 1 samples (green circles, n = 240); significant SNPs in combined dataset of stage 2 (red diamond, n = 1,156). This region has *RGS7* genes related to one significant SNP.



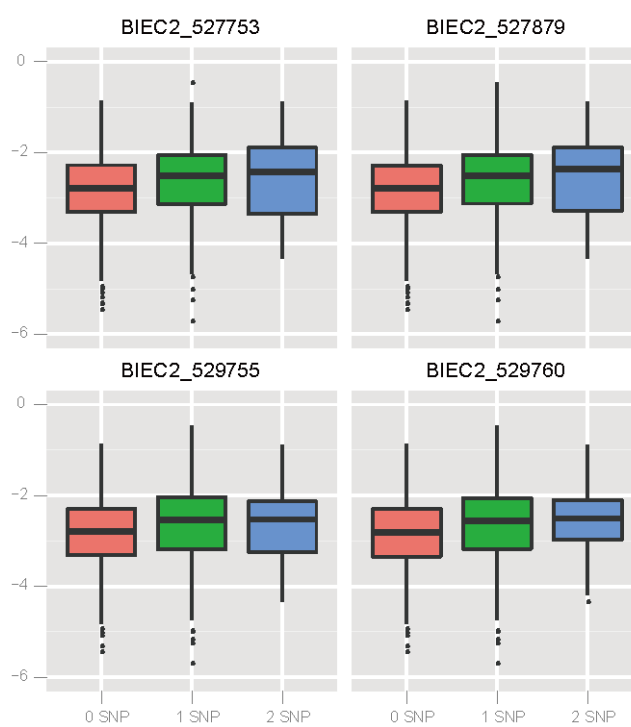
Supplementary Figure S7

Boxplots show cumulative effect for EBV of the effect allele number of significant 6 SNPs on chromosome 21.



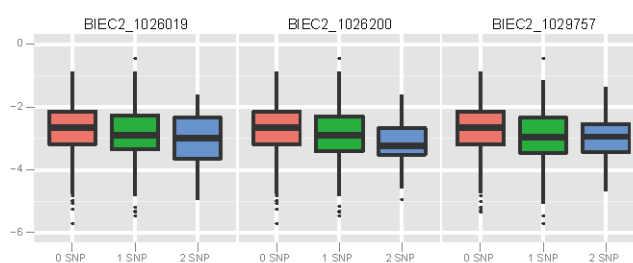
Supplementary Figure S8

Boxplots show cumulative effect for EBV of the effect allele number of significant 4 SNPs on chromosome 20.



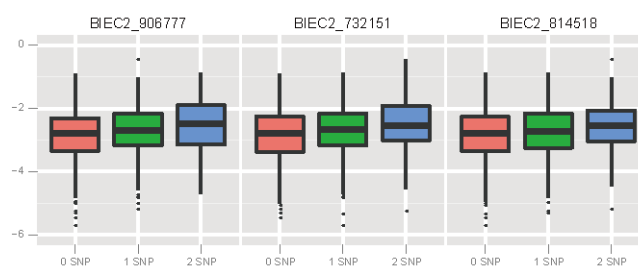
Supplementary Figure S9

Boxplots show cumulative effect for EBV of the effect allele number of significant 3 SNPs on chromosome 8.



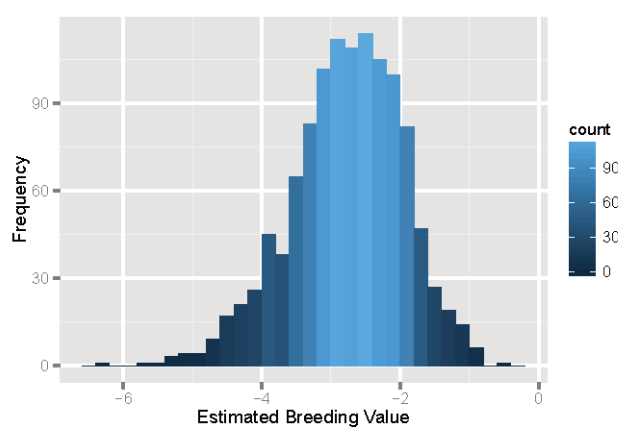
Supplementary Figure S10

Boxplots show cumulative effect for EBV of the effect allele number of significant each one SNPs on chromosome 5, 28, 30.



Supplementary Figure S11

Histogram of EBVs for 1,156 Thoroughbreds.



Supplementary Table S1

List of two reported SNPs that associated with racing performance of Thoroughbreds.

	Chr	BP	Neareat Gene (Ensemble Gene ID)	SNP Type	Minor Allele	Major Allele	Stage 1			Stage 2		
							P-value	MAF	SNP effect	P-value	MAF	SNP effect
BIEC2_417274	18	65868604	-	InterGenic	C	A	0.801	0.498	-0.02	5.38.E-04	0.491	-0.124
BIEC2_417495	18	67186093	-	InterGenic	G	A	0.526	0.483	-0.05	4.97.E-04	0.480	-0.123