

# A genome-wide association study for body weight in Japanese Thoroughbred racehorses clarifies candidate regions on chromosomes 3, 9, 15, and 18

Teruaki TOZAKI<sup>1\*</sup>, Mio KIKUCHI<sup>1</sup>, Hironaga KAKOI<sup>1</sup>, Kei-ichi HIROTA<sup>1</sup> and Shun-ichi NAGATA<sup>1</sup>

<sup>1</sup>Genetic Analysis Department, Laboratory of Racing Chemistry, Tochigi 320-0851, Japan

*Body weight is an important trait to confirm growth and development in humans and animals. In Thoroughbred racehorses, it is measured in the postnatal, training, and racing periods to evaluate growth and training degrees. The body weight of mature Thoroughbred racehorses generally ranges from 400 to 600 kg, and this broad range is likely influenced by environmental and genetic factors. Therefore, a genome-wide association study (GWAS) using the Equine SNP70 BeadChip was performed to identify the genomic regions associated with body weight in Japanese Thoroughbred racehorses using 851 individuals. The average body weight of these horses was 473.9 kg (standard deviation: 28.0) at the age of 3, and GWAS identified statistically significant SNPs on chromosomes 3 (BIEC2\_808466,  $P=2.32E-14$ ), 9 (BIEC2\_1105503,  $P=1.03E-7$ ), 15 (BIEC2\_322669,  $P=9.50E-6$ ), and 18 (BIEC2\_417274,  $P=1.44E-14$ ), which were associated with body weight as a quantitative trait. The genomic regions on chromosomes 3, 9, 15, and 18 included ligand-dependent nuclear receptor compressor-like protein (LCORL), zinc finger and AT hook domain containing (ZFAT), tribbles pseudokinase 2 (TRIB2), and myostatin (MSTN), respectively, as candidate genes. LCORL and ZFAT are associated with withers height in horses, whereas MSTN affects muscle mass. Thus, the genomic regions identified in this study seem to affect the body weight of Thoroughbred racehorses. Although this information is useful for breeding and growth management of the horses, the production of genetically modified animals and gene doping (abuse/misuse of gene therapy) should be prohibited to maintain horse racing integrity.*

**Key words:** body weight, LCORL, MSTN, Thoroughbred, ZFAT

J. Equine Sci.  
Vol. 28, No. 4  
pp. 127–134, 2017

The Thoroughbred was developed in the early 18th century, and all modern Thoroughbred racehorses can be traced back to a small number of Arabian, Barb, and Turkmen stallions and many native British mares living approximately 300 years ago [1, 4, 31]. At present, about 100,000 foals are born worldwide every year, and almost all are used as racehorses.

Thoroughbred racehorses have been bred for improved

racing ability, such as speed and stamina, which has led to variations in morphological traits such as body weight. This characteristic is used for confirming growth in humans and animals, and in Thoroughbred racehorses, it is measured at several occasions including postnatal, training, and racing periods to confirm growth and training degrees [13–16].

The body weight of mature Thoroughbred racehorses generally ranges from 400 to 600 kg. In Japanese Thoroughbred racehorses, the average body weight at a horse's first start (897 days) is 455 kg [10], whereas the average body weights of Korean Thoroughbreds are 460, 454, and 441 kg for stallions, geldings, and mares, respectively [2]. Several genetic and environmental factors, such as sire, sex, age, and facility, affect body weight in this breed [10].

Heritability estimates of body weight in no-defined-breed horses belonging to the Brazilian army cavalry and

Received: July 3, 2017

Accepted: August 21, 2017

\*Corresponding author. e-mail: ttozaki@lrc.or.jp

©2017 Japanese Society of Equine Science

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

in Japanese Thoroughbred racehorses were  $0.40 \pm 0.034$  and  $0.27 \pm 0.04$ , respectively [10, 22]. In Korean Thoroughbred racehorses, heritability estimates of body weight were 0.578, 0.472, and 0.555 for stallions, geldings, and mares, respectively [2]. Thus, genetic factors appear to contribute to the individual differences in body weight in Thoroughbred racehorses. Recently, horse genome mapping has significantly advanced, and with the completion of the Horse Genome Project [29], a single nucleotide polymorphisms (SNPs) map containing more than one million markers is now available. This allows a genome-wide association study (GWAS) that facilitates the discovery of genetic traits that are of interest to breeders, trainers, and owners.

Although body weight is a highly complex trait and its genetic basis is not well elucidated in Thoroughbred racehorses, several genetic factors might contribute to differences among individuals. Therefore, a GWAS using the Equine SNP70 BeadChip (65,156 SNPs) was designed and performed in the present study, aiming to identify the genomic regions associated with body weight in Japanese Thoroughbred racehorses.

## Materials and Methods

### *Body weight of Thoroughbred racehorses*

This study was performed according to the ethical standards in animal research. Body weight (kg) is measured when each Thoroughbred horse attends a race in the Japan Racing Association (JRA). The body weights in the final race of each horse at the ages of 2, 3, and 4 years were collected from official JRA race data, and most measurements were taken from September to December. The complete body weight records included 535 horses at the age of 2, 851 horses at the age of 3, and 734 horses at the age of 4 years. In our GWAS, we used only the body weights of 3-year-old horses because more phenotypic information was available for these horses compared with horses of other ages. The 851 three-year-old Thoroughbred racehorses (619 males and 232 females) used in the GWAS were well trained as racehorses and experienced in racing in the JRA.

### *Genome-wide association study*

Blood samples were collected from each animal and stored at  $-40^{\circ}\text{C}$ . Genomic DNA was extracted using an MFX-2000 MagExtractor System (Toyobo, Osaka, Japan), according to the manufacturer's protocol. Almost all blood samples were also used in previous studies [23, 24]. The quantity and purity of the extracted DNA were assessed by spectrophotometric absorbance, and the 260/230 and 260/280 ratios were validated.

Samples were genotyped using the EquineSNP70 Genotyping BeadChips (Illumina, San Diego, CA, U.S.A.). This

array contains 65,157 SNPs ascertained from the EquCab2 database of the horse genome [29]. Genotyping was performed at Neogen GeneSeek (Lincoln, NE, U.S.A.). All samples had a genotyping rate above 90%. We omitted SNPs that had a genotyping completion rate below 99%, deviated from Hardy-Weinberg equilibrium (HWE,  $P < 1.0 \times 10^{-6}$ ), or were monomorphic or had minor allele frequencies (MAFs) below 5% in our samples. Thus, 44,306 SNPs were used in this study.

GWAS was performed between the 44,306 SNPs and body weight as a quantitative trait of the 851 horses (619 males and 232 females). All statistical analyses for GWAS including quality control were performed using PLINK version 1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) [18]. Manhattan and Q-Q plots were generated with the qqman R package [28].

### *Statistical analysis*

The mean, standard deviation (SD), quartiles, minimum, and maximum body weights at each age were calculated. The differences in body weight among ages were evaluated by analysis of variance (ANOVA). The differences in body weight among genotypes of each candidate gene were evaluated by ANOVA.

The proportion of weight variation explained was estimated using a normal linear model and by comparing the residual variance of a null model with sex only (VN) to a full model (VF) with sex and relevant markers. The proportion of explained variance is defined as  $1 - (VF/VN)$ .

All statistical analyses were performed using R (<http://www.R-project.org/>).

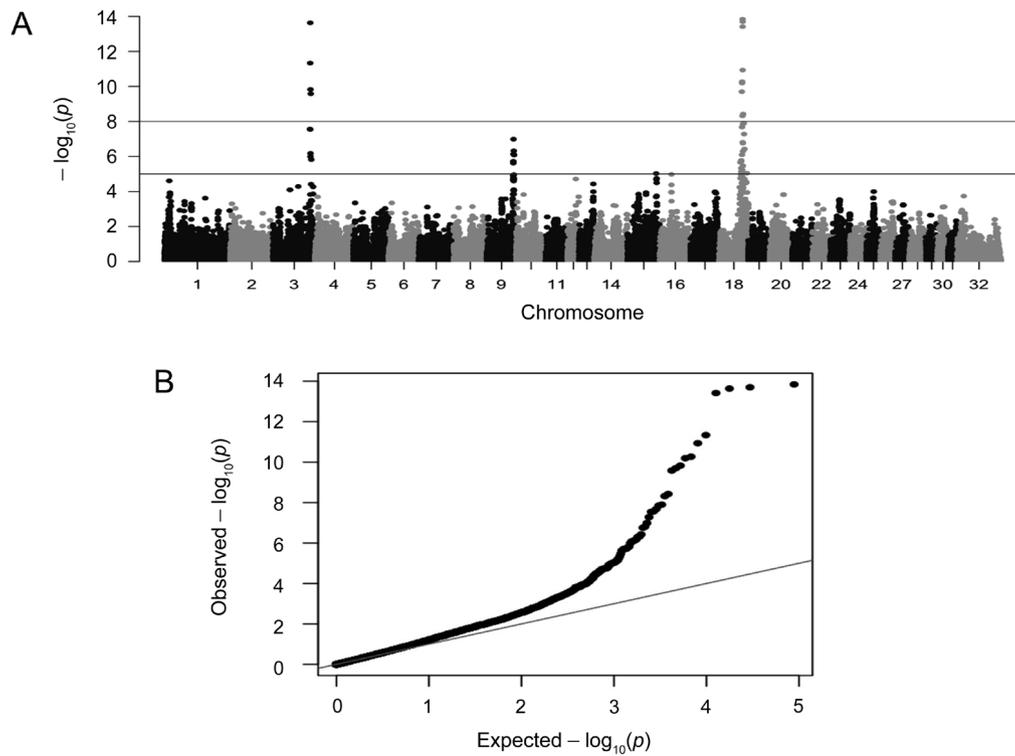
## Results

### *Body weight variation*

The average body weights and standard deviations of the 2-, 3-, and 4-year-old horses (535, 851, and 734 individuals, respectively) were  $468.8 \pm 26.1$ ,  $473.9 \pm 28.0$ , and  $478.8 \pm 27.6$  kg, respectively (Table 1), and a statistically significant

**Table 1.** Body weight of Japanese Thoroughbred racehorses of three different ages

	2 years old	3 years old	4 years old
Sample number (N)	535	851	734
Mean (kg)	468.8	473.9	478.8
SD (kg)	26.1	28.0	27.6
Minimum (kg)	382	390	395
First quartile (kg)	452	456	460
Second quartile (kg)	468	473	478
Third quartile (kg)	485	492	496
Maximum (kg)	568	568	604



**Fig. 1.** Genome-wide association study (GWAS) for body weight in Japanese Thoroughbred racehorses. A) Manhattan plot of the  $-\log_{10}(P)$  from a GWAS of body weight in Japanese Thoroughbreds. B) Q-Q plot of observed versus expected  $-\log_{10}(P)$  from a GWAS in Japanese Thoroughbred racehorses.

difference was observed between them (ANOVA:  $P < 0.01$ ). Thus, body weight increased with age.

Minimum, quartiles, and maximum body weights showed trends similar to that of average body weight. The minimum weight observed in the present study was 382 kg in a 2-year-old racehorse, and the maximum was 604 kg in a 4-year-old racehorse, which had an extreme weight gain (36 kg) at this age.

#### Genome-wide association study

The results from GWAS are represented graphically in a Manhattan plot (Fig. 1a) and a Q-Q plot (Fig. 1b). The nine SNPs located at 104,403,770–108,176,636 bp (spanning 3.77 Mb) on chromosome 3, seven SNPs located at 74,756,685–77,100,429 bp (spanning 2.34 Mb) on chromosome 9, one SNP located at 80,921,871 bp on chromosome 15, and 30 SNPs located at 59,614,257–78,131,099 bp (spanning 18.5 Mb) on chromosome 18 were associated with body weight as a qualitative trait ( $P < 1.00E-8$  and/or  $P < 1.00E-5$ ) in our GWAS (Table 2). The SNPs *BIEC2\_808466* ( $P = 2.32E-14$ ) on chromosome 3, *BIEC2\_1105503* ( $P = 1.03E-07$ ) on chromosome 9, *BIEC2\_322669* ( $P = 9.50E-06$ ) on chromosome 15, and *BIEC2\_417274* on chromosome 18 ( $P = 1.44E-14$ ) revealed the most statistically significant differences for

body weight at each genomic region in our GWAS (Table 2).

#### Genes annotated on each chromosome

Nineteen protein-coding genes, seven noncoding RNAs (ncRNAs), and one pseudogene have been annotated in the 104,403,770–108,176,636 bp (spanning 3.77 Mb) region of chromosome 3 (Table 3). This genomic region contains the *ligand-dependent nuclear receptor compressor-like protein (LCORL)* gene.

Two protein-coding genes, two ncRNAs, and two microRNAs have been annotated in the 74,756,685–77,100,429 bp (spanning 2.34 Mb) region of chromosome 9 (Table 3). This genomic region contains the *zinc finger and AT hook domain containing (ZFAT)* gene.

In the 59,614,257–78,131,099 bp (spanning 18.52 Mb) region of chromosome 18, 108 protein-coding genes, 11 ncRNAs, one microRNA, and 12 pseudogenes have been annotated (Table 3). This genomic region contains the *myostatin (MSTN)* gene. In addition, *tribbles pseudokinase 2 (TRIB2)*, a protein-coding gene, has been annotated near *BIEC2\_322669* (80,921,871 bp) on chromosome 15 (Table 3).

#### The effect of genotypic differences

Table 4 shows the average body weight of each

**Table 2.** Results of the genome-wide association study conducted for the body weight of Japanese Thoroughbred racehorses using PLINK

Chromosome	SNP	Position (bp)	BETA	SE	R <sup>2</sup>	T	P
3	<i>BIEC2_808344</i>	104,403,770	8.639	1.539	0.03577	5.612	2.71E-08
3	<i>BIEC2_808426</i>	104,941,985	12.180	1.735	0.05489	7.018	4.62E-12
3	<i>BIEC2_808432</i>	104,986,284	9.279	1.656	0.03572	5.601	2.87E-08
3	<i>BIEC2_808466</i>	105,163,077	17.140	2.206	0.06633	7.766	2.32E-14
3	<i>BIEC2_808608</i>	105,875,809	-6.841	1.389	0.02783	-4.924	1.02E-06
3	<i>BIEC2_808617</i>	105,876,397	8.985	1.385	0.04728	6.487	1.49E-10
3	<i>BIEC2_808640</i>	105,947,243	7.933	1.584	0.02873	5.009	6.67E-07
3	<i>BIEC2_808825</i>	106,710,385	9.541	1.491	0.04599	6.397	2.61E-10
3	<i>BIEC2_809156</i>	108,176,636	6.923	1.429	0.02695	4.844	1.51E-06
9	<i>BIEC2_1165712</i>	74,756,685	7.002	1.474	0.02590	4.749	2.40E-06
9	<i>BIEC2_1105370</i>	74,795,013	6.905	1.440	0.02639	4.794	1.93E-06
9	<i>BIEC2_1105377</i>	74,798,143	6.905	1.440	0.02639	4.794	1.93E-06
9	<i>BIEC2_1165849</i>	75,245,697	7.284	1.460	0.02853	4.987	7.43E-07
9	<i>BIEC2_1105503</i>	75,374,649	7.975	1.486	0.03282	5.368	1.03E-07
9	<i>BIEC2_1105810</i>	76,167,910	6.908	1.361	0.02944	5.075	4.76E-07
9	<i>BIEC2_1105995</i>	77,100,429	7.948	1.601	0.02826	4.966	8.27E-07
15	<i>BIEC2_322669</i>	80,921,871	8.871	1.991	0.02303	4.455	9.50E-06
18	<i>BIEC2_416419</i>	59,614,257	8.170	1.713	0.02634	4.770	2.17E-06
18	<i>BIEC2_416680</i>	62,049,912	7.233	1.500	0.02682	4.822	1.68E-06
18	<i>BIEC2_416681</i>	62,054,146	6.872	1.491	0.02444	4.609	4.67E-06
18	<i>BIEC2_416683</i>	62,115,222	6.724	1.490	0.02351	4.513	7.28E-06
18	<i>BIEC2_416689</i>	62,117,055	6.686	1.504	0.02277	4.445	9.94E-06
18	<i>BIEC2_416704</i>	62,148,769	7.220	1.505	0.02654	4.797	1.90E-06
18	<i>BIEC2_438369</i>	63,193,752	-6.071	1.345	0.02344	-4.514	7.24E-06
18	<i>BIEC2_416921</i>	63,708,318	8.945	1.389	0.04701	6.441	1.99E-10
18	<i>BIEC2_416992</i>	64,245,075	-6.740	1.356	0.02836	-4.972	8.03E-07
18	<i>BIEC2_438541</i>	64,252,426	7.601	1.344	0.03645	5.657	2.10E-08
18	<i>BIEC2_417010</i>	64,391,698	9.084	1.372	0.04930	6.620	6.40E-11
18	<i>BIEC2_417030</i>	64,481,405	6.039	1.347	0.02320	4.482	8.40E-06
18	<i>BIEC2_417075</i>	64,725,066	8.965	1.349	0.04946	6.646	5.37E-11
18	<i>BIEC2_417120</i>	64,919,213	-6.439	1.412	0.02399	-4.560	5.86E-06
18	<i>BIEC2_417187</i>	65,565,128	-7.881	1.332	0.03963	-5.915	4.80E-09
18	<i>BIEC2_417274</i>	65,868,604	10.150	1.297	0.06737	7.831	1.44E-14
18	<i>UKUL3221</i>	65,969,033	9.934	1.290	0.06531	7.698	3.85E-14
18	<i>BIEC2_417291</i>	66,010,474	-8.987	1.306	0.05281	-6.880	1.16E-11
18	<i>BIEC2_417308</i>	66,155,365	-7.025	1.328	0.03211	-5.291	1.55E-07
18	<i>BIEC2_438865</i>	66,158,121	10.090	1.295	0.06666	7.787	2.00E-14
18	<i>BIEC2_417372</i>	66,539,967	-7.253	1.377	0.03180	-5.268	1.75E-07
18	<i>BIEC2_417423</i>	66,819,091	-7.847	1.370	0.03726	-5.729	1.40E-08
18	<i>BIEC2_438994</i>	66,862,437	-6.296	1.412	0.02289	-4.460	9.29E-06
18	<i>BIEC2_438997</i>	66,892,380	-6.771	1.449	0.02508	-4.673	3.45E-06
18	<i>BIEC2_417454</i>	66,996,889	7.087	1.399	0.02951	-5.066	5.00E-07
18	<i>BIEC2_417524</i>	67,545,703	-8.047	1.351	0.04012	-5.957	3.76E-09
18	<i>BIEC2_417704</i>	69,258,120	-7.792	1.356	0.03750	-5.748	1.26E-08
18	<i>BIEC2_417806</i>	69,969,173	-7.272	1.324	0.03468	-5.493	5.24E-08
18	<i>BIEC2_418057</i>	71,300,017	-7.885	1.540	0.02997	-5.121	3.75E-07
18	<i>BIEC2_421048</i>	78,131,099	-6.230	1.395	0.02297	-4.465	9.11E-06

SNP, single nucleotide polymorphism; BETA, regression coefficient; SE, standard error; R<sup>2</sup>, regression r-squared; T, wald test (based on t-distribution); P, wald test asymptotic P-value.

genotype characterized by the most statistically different SNPs. We observed a statistically significant difference in body weight depending on genotypes at *BIEC2\_808466*, *BIEC2\_1105503*, *BIEC2\_322669*, and *BIEC2\_417274*.

The four SNPs on chromosomes 3, 9, 15, and 18 explained 17.4% of the weight variance observed in the 851 three-year-old Japanese Thoroughbred racehorses. These four loci together with gender explained 30.0% of the weight

**Table 3.** Number of annotated genes in the candidate genomic regions on chromosomes 3, 9, 15, and 18

	Chromosome 3 (3.77 Mb)	Chromosome 9 (2.34 Mb)	Chromosome 15	Chromosome 18 (18.52 Mb)
Protein coding gene	19	2	1	108
Noncoding RNA	7	2	0	11
MicroRNA	0	2	0	1
Pseudogene	1	0	0	12
Total	27	6	1	132

**Table 4.** Body weight of each genotype defined by the single nucleotide polymorphisms found close to candidate genes in each chromosome

Chromosome	SNP	Position (bp)	MAF	P-value	Candidate gene	Homozygote (kg)	Heterozygote (kg)	Homozygote (kg)	ANOVA
3	<i>BIEC2_808466</i>	105,163,077	0.112	2.32E-14	<i>LCORL</i>	470.1	487.1	510.0	0.00024
9	<i>BIEC2_1105503</i>	75,374,649	0.273	1.03E-07	<i>ZFAT</i>	470.0	476.4	488.4	<1.0E-4
15	<i>BIEC2_322669</i>	80,921,871	0.133	9.50E-06	<i>TRIB2</i>	471.7	480.4	490.2	<1.0E-4
18	<i>BIEC2_417274</i>	66,158,121	0.478	1.44E-14	<i>MSTN</i>	457.5	479.4	482.1	<1.0E-4

SNP, single nucleotide polymorphism; MAF, minor allele frequency; ANOVA, analysis of variance.

variance observed in the Japanese Thoroughbred racehorses.

## Discussion

Our GWAS using 851 Japanese Thoroughbred racehorses identified the regions near *LCORL* on chromosome 3, *ZFAT* on chromosome 9, *TRIB2* on chromosome 15, and *MSTN* on chromosome 18 as candidate genes for body weights. Similar results were also observed in GWAS performed using body weight information of 2-year-olds (535 horses) and 4-year-olds (734 horses) (data not shown). Therefore, these genomic regions are good candidates for regions associated with body weight as these four loci explained 17.4% of body weight variance in this horse breed. Because the heritability estimate of body weight at a horse's first start was  $0.27 \pm 0.04$  in Japanese Thoroughbreds [10], these four loci were expected to explain much of the genetic variance in these horses.

In this study, the body weights used were from between September and December of each year. Small differences in age due to birth month and measurement month may influence body weight. However, because Thoroughbred horses are nearly mature by 3 years of age, this influence was excluded from the model in this study.

*LCORL* encodes a transcription factor that appears to function in spermatogenesis and has been reported to affect withers height in horses. It has been reported that the most significant SNP to withers height is *BIEC2\_808543* (105,547,002 bp) [8, 9, 19], which is located 63.3 kb upstream of *LCORL*. *BIEC2\_808466* (105,163,077 bp), which was identified in the present study, is located 447.3 kb upstream of *LCORL*. *BIEC2\_808543* was not identified

in the present study due to the quality control performed in our GWAS, which excluded *BIEC2\_808543*. Because no coding genes were found in the genomic region between *BIEC2\_808466* and *LCORL*, it is considered that our GWAS identified *LCORL* as a candidate gene for body weight in Japanese Thoroughbred racehorses.

In a previous study, *BIEC2\_808543* near *LCORL* was associated with withers height and cannon circumference in Japanese Thoroughbred horses under training (September to April; 1- to 2-year-olds), which affected their body weights [26]. This could also apply to mature horses. Because *LCORL* affects skeletal frame and body size in humans and animals [6, 20], changes in the skeletal frame are considered to be associated with body size, reflecting body weight. These findings also suggest that *LCORL* is the best candidate for body weight in the genomic region of chromosome 3.

*ZFAT* encodes a protein that likely binds DNA and functions as a transcriptional regulator involved in apoptosis and cell survival, and it has an essential role in hematopoietic differentiation in blood islands [27]. In a previous GWAS, a region from 74,795,013 to 76,254,733 bp including *ZFAT* was identified as a candidate region for withers height in horses [8]. *BIEC2\_1105370* and *BIEC2\_1105377*, which were identified in the present study, were also detected in that previous study. Because *ZFAT* is associated with body height in Japanese and Korean human populations [21] and withers height in several horse breeds [8, 9, 19], it is expected that differences in withers height due to *ZFAT* affects body weight in a similar fashion as that reported for *LCORL*. However, because *ZFAT* and/or its associated SNPs contribute less than *LCORL* to withers height in Thoroughbred racehorses [8], further detailed analyses are required.

*TRIB2*, one of the three members of the Tribbles family, encodes a scaffold protein involved in intracellular signal transduction [5], and, interestingly, it is associated with the internal fat area and volume of the pericardial fat [12]. Although there are no reports on the association between morphological traits and fat metabolism in horses, and it is still unclear if internal fat area and volume of pericardial fat affect body weight; it is possible that may contribute to it. While only *BIEC2\_322669* was identified and located 178.9 kb downstream of *TRIB2*, *BIEC2\_323180* ( $P=2.70E-5$ ) was located 350.3 kb upstream of this gene. Unfortunately, the 10 SNPs between 80,950,328 and 81,428,857 bp, which is the region where *TRIB2* is located, had no statistical significance.

*MSTN* is a member of the TGF-beta (transforming growth factor-beta) superfamily of proteins. This protein negatively regulates skeletal muscle cell proliferation and differentiation. Mutations in this gene are associated with increased skeletal muscle mass in humans and other mammals. In addition, *MSTN* is associated with optimal race distance in Thoroughbreds by affecting the ratio of muscle fiber types I, IIa, and IIx. Three SNPs, *BIEC2\_417274*, *UKUL3221*, and *BIEC2\_438865*, were the most significant in this candidate gene region, and *BIEC2\_438865* was located 332 kb downstream of *MSTN*. In addition, *BIEC2\_417274* has been identified as a candidate SNP for racing performance and optimal race distance in Japanese Thoroughbred racehorses [23].

In a previous study, our group also showed that several SNPs near *MSTN* were associated with muscle mass at withers height in Japanese Thoroughbreds under training (September to April; 1- to 2-year-olds), which also affected body weight [25]. This could also apply to mature horses. Because *MSTN* affects the muscle mass of animals [11, 17], it is considered that changes in muscle mass may affect body weight. These findings suggest that *MSTN* is a good candidate gene for body weight on chromosome 18.

The four genes *LCORL*, *ZFAT*, *TRIB2*, and *MSTN* affected Thoroughbred body weight, which ranged from 470.1 to 510.0 kg, 470.0 to 488.4 kg, 471.7 to 490.2 kg, and 457.5 to 482.1 kg, respectively (Table 4), and *LCORL* is expected to contribute to the largest weight changes. Because the minor allele frequency (MAF) for *LCORL* was 0.0125 (Table 4), selective breeding to increase less-frequent genotypes will contribute to higher weights in Thoroughbreds.

*High-mobility group AT-hook 2 (HMGA2)* on chromosome 6 and *LIM and SH3 domain protein 1 (LASP1)* on chromosome 11 were not identified as impacting body weight in our study, although these genes, in addition to *LCORL* and *ZFAT*, are candidates for withers height or size variation in horses [19]. The architectural transcription factor *HMGA2* regulates gene expression and directs

cellular growth, proliferation, and differentiation [3], and it is associated with human height [30]. Knockout of this gene in mice resulted in individuals with only 40% of the body weight of control individuals [32]. *LASP1* mediates cell migration and survival, and its expression is induced by insulin-like growth factor-1 (IGF1) [7]; its misexpression in mice disrupts chondrocyte differentiation. Therefore, *HMGA2* and *LASP1* were also expected candidate genes for body weight in Thoroughbred racehorses. However, because they were not identified in our GWAS, it is considered that variants of these genes are not present in Thoroughbreds or that these genes contribute less to Thoroughbred body weight than *LCORL* and *ZFAT*.

Although the genes and SNPs identified in this study are useful for effective feeding management, training, and breeding, there is also concern about their use for the production of genetically modified racehorses or possible gene doping in mature individuals. For the purpose of fair horse racing enforcement, the International Federation of Horseracing Authorities (IFHA) prohibits genetic modification and gene doping. Therefore, misuse of these genes needs to be monitored, or an inspection method for gene doping needed to be created in order to maintain horse racing integrity.

## Conflict of Interest

The Laboratory of Racing Chemistry (LRC) is licensed by Plusvital (NovaUCD, Belfield, Dublin, Ireland) to perform the Speed Gene Test for genotyping *MSTN* (Japanese Patent No. 5667057) and provides genotyping services to clients in Japan. The LRC also conducts the withers height test in Thoroughbred racehorses by genotyping *LCORL*.

## Acknowledgments

We would like to thank Dr. M. Kurosawa for useful discussions and Ms. H. Sato for her helpful assistance during this study. We would like to thank the JRA Equine Department for approving and supporting this study through a grant-in-aid (2014–2016).

## References

1. Bower, M.A., Campana, M.G., Whitten, M., Edwards, C.J., Jones, H., Barrett, E., Cassidy, R., Nisbet, R.E., Hill, E.W., Howe, C.J., and Binns, M. 2011. The cosmopolitan maternal heritage of the Thoroughbred racehorse breed shows a significant contribution from British and Irish native mares. *Biol. Lett.* 7: 316–320. [Medline] [CrossRef]
2. Cho, K.H., Son, S.K., Cho, B.W., Lee, H.K., Kong, H.S., Jeon, G.J., and Park, K.D. 2008. Effects of change of body weight on racing time in thoroughbred racehorses. *J. Anim.*

- Sci. Technol. Korea* **50**: 741–746. [[CrossRef](#)]
3. Cleynen, I., and Van de Ven, W.J. 2008. The HMGA proteins: a myriad of functions (Review). *Int. J. Oncol.* **32**: 289–305. [[Medline](#)]
  4. Cunningham, E.P., Dooley, J.J., Splan, R.K., and Bradley, D.G. 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Anim. Genet.* **32**: 360–364. [[Medline](#)] [[CrossRef](#)]
  5. Grandinetti, K.B., Stevens, T.A., Ha, S., Salamone, R.J., Walker, J.R., Zhang, J., Agarwalla, S., Tenen, D.G., Peters, E.C., and Reddy, V.A. 2011. Overexpression of TRIB2 in human lung cancers contributes to tumorigenesis through downregulation of C/EBP $\alpha$ . *Oncogene* **30**: 3328–3335. [[Medline](#)] [[CrossRef](#)]
  6. Lindholm-Perry, A.K., Sexten, A.K., Kuehn, L.A., Smith, T.P., King, D.A., Shackelford, S.D., Wheeler, T.L., Ferrell, C.L., Jenkins, T.G., Snelling, W.M., and Freetly, H.C. 2011. 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.* **12**: 103. [[Medline](#)] [[CrossRef](#)]
  7. Loughran, G., Huigsloot, M., Kiely, P.A., Smith, L.M., Floyd, S., Ayllon, V., and O'Connor, R. 2005. Gene expression profiles in cells transformed by overexpression of the IGF-I receptor. *Oncogene* **24**: 6185–6193. [[Medline](#)] [[CrossRef](#)]
  8. Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Loreda, A.I., Bellone, R.R., Mezey, J.G., Brooks, S.A., and Sutter, N.B. 2012. Four loci explain 83% of size variation in the horse. *PLoS One* **7**: e39929. [[Medline](#)] [[CrossRef](#)]
  9. Metzger, J., Schrimpf, R., Philipp, U., and Distl, O. 2013. Expression levels of LCORL are associated with body size in horses. *PLoS One* **8**: e56497. [[Medline](#)] [[CrossRef](#)]
  10. Moritsu, Y., Deki, T.T., Miwa, K., and Ichikawa, S. 1997. Estimates of heritability and environmental factors in age and body weight at first start among 3-year-old Thoroughbred horses. *Anim. Sci. Agric. Hokkaido* **39**: 15–20 (in Japanese).
  11. Mosher, D.S., Quignon, P., Bustamante, C.D., Sutter, N.B., Mellersh, C.S., Parker, H.G., and Ostrander, E.A. 2007. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet.* **3**: e79. [[Medline](#)] [[CrossRef](#)]
  12. Nakayama, K., Ogawa, A., Miyashita, H., Tabara, Y., Igase, M., Kohara, K., Miki, T., Kagawa, Y., Yanagisawa, Y., Katashima, M., Onda, T., Okada, K., Fukushima, S., and Iwamoto, S. 2013. Positive natural selection of TRIB2, a novel gene that influences visceral fat accumulation, in East Asia. *Hum. Genet.* **132**: 201–217. [[Medline](#)] [[CrossRef](#)]
  13. Onoda, T., Yamamoto, R., Sawamura, K., Inoue, Y., Matsui, A., Miyake, T., and Hirai, N. 2011. Empirical growth curve estimation using sigmoid sub-functions that adjust seasonal compensatory growth for male body weight of Thoroughbred horses. *J. Equine Sci.* **22**: 37–42. [[Medline](#)] [[CrossRef](#)]
  14. Onoda, T., Yamamoto, R., Sawamura, K., Inoue, Y., Murase, H., Nambo, Y., Tozaki, T., Matsui, A., Miyake, T., and Hirai, N. 2013. Empirical growth curve estimation considering multiple seasonal compensatory growths of body weights in Japanese Thoroughbred colts and fillies. *J. Anim. Sci.* **91**: 5599–5604. [[Medline](#)] [[CrossRef](#)]
  15. Onoda, T., Yamamoto, R., Sawamura, K., Murase, H., Nambo, Y., Inoue, Y., Matsui, A., Miyake, T., and Hirai, N. 2013. Empirical percentile growth curves with z-scores considering seasonal compensatory growths for Japanese Thoroughbred horses. *J. Equine Sci.* **24**: 63–69. [[Medline](#)] [[CrossRef](#)]
  16. Onoda, T., Yamamoto, R., Sawamura, K., Murase, H., Nambo, Y., Inoue, Y., Matsui, A., Miyake, T., and Hirai, N. 2014. An approach of estimating individual growth curves for young thoroughbred horses based on their birthdays. *J. Equine Sci.* **25**: 29–35. [[Medline](#)] [[CrossRef](#)]
  17. O'Rourke, B.A., Greenwood, P.L., Arthur, P.F., and Goddard, M.E. 2013. Inferring the recent ancestry of myostatin alleles affecting muscle mass in cattle. *Anim. Genet.* **44**: 86–90. [[Medline](#)] [[CrossRef](#)]
  18. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**: 559–575. [[Medline](#)] [[CrossRef](#)]
  19. Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T., and Rieder, S. 2012. A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS One* **7**: e37282. [[Medline](#)] [[CrossRef](#)]
  20. Soranzo, N., Rivadeneira, F., Chinappan-Horsley, U., Malkina, I., Richards, J.B., Hammond, N., Stolk, L., Nica, A., Inouye, M., Hofman, A., Stephens, J., Wheeler, E., Arp, P., Gwilliam, R., Jhamai, P.M., Potter, S., Chaney, A., Ghori, M.J., Ravindrarajah, R., Ermakov, S., Estrada, K., Pols, H.A., Williams, F.M., McArdle, W.L., van Meurs, J.B., Loos, R.J., Dermitzakis, E.T., Ahmadi, K.R., Hart, D.J., Ouwehand, W.H., Wareham, N.J., Barroso, I., Sandhu, M.S., Strachan, D.P., Livshits, G., Spector, T.D., Uitterlinden, A.G., and Deloukas, P. 2009. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet.* **5**: e1000445. [[Medline](#)] [[CrossRef](#)]
  21. Takeuchi, F., Nabika, T., Isono, M., Katsuya, T., Sugiyama, T., Yamaguchi, S., Kobayashi, S., Yamori, Y., Ogihara, T., and Kato, N. 2009. Evaluation of genetic loci influencing adult height in the Japanese population. *J. Hum. Genet.* **54**: 749–752. [[Medline](#)] [[CrossRef](#)]
  22. Tamioso, P.R., Cosmo, T.R., Pimentel, C.M.M., Dias, L.T.,

- and Teixeira, R.A. 2012. Heritability estimates for body weight and height at withers in Brazilian army horses. *Cienc. Rural. Santa Maria* **42**: 2246–2251. [[CrossRef](#)]
23. Tozaki, T., Hill, E.W., Hirota, K., Kakoi, H., Gawahara, H., Miyake, T., Sugita, S., Hasegawa, T., Ishida, N., Nakano, Y., and Kurosawa, M. 2012. A cohort study of racing performance in Japanese Thoroughbred racehorses using genome information on ECA18. *Anim. Genet.* **43**: 42–52. [[Medline](#)] [[CrossRef](#)]
  24. Tozaki, T., Miyake, T., Kakoi, H., Gawahara, H., Sugita, S., Hasegawa, T., Ishida, N., Hirota, K., and Nakano, Y. 2010. A genome-wide association study for racing performances in Thoroughbreds clarifies a candidate region near the MSTN gene. *Anim. Genet.* **41**(Suppl 2): 28–35. [[Medline](#)] [[CrossRef](#)]
  25. Tozaki, T., Sato, F., Hill, E.W., Miyake, T., Endo, Y., Kakoi, H., Gawahara, H., Hirota, K., Nakano, Y., Nambo, Y., and Kurosawa, M. 2011. Sequence variants at the myostatin gene locus influence the body composition of Thoroughbred horses. *J. Vet. Med. Sci.* **73**: 1617–1624. [[Medline](#)] [[CrossRef](#)]
  26. Tozaki, T., Sato, F., Ishimaru, M., Kikuchi, M., Kakoi, H., Hirota, K.I., and Nagata, S.I. 2016. Sequence variants of BIEC2-808543 near LCORL are associated with body composition in Thoroughbreds under training. *J. Equine Sci.* **27**: 107–114. [[Medline](#)] [[CrossRef](#)]
  27. Tsunoda, T., Takashima, Y., Tanaka, Y., Fujimoto, T., Doi, K., Hirose, Y., Koyanagi, M., Yoshida, Y., Okamura, T., Kuroki, M., Sasazuki, T., and Shirasawa, S. 2010. Immune-related zinc finger gene ZFAT is an essential transcriptional regulator for hematopoietic differentiation in blood islands. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 14199–14204. [[Medline](#)] [[CrossRef](#)]
  28. Turner, S.D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* doi: 10.1101/005165.
  29. Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerrer, S., Inslan, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., Blöcker, H., Distl, O., Edgar, R.C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J.N., Penedo, M.C., Raison, J.M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D.F., Biagi, T., Binns, M.M., Chowdhary, B.P., Coleman, S.J., Della Valle, G., Fryc, S., Guérin, G., Hasegawa, T., Hill, E.W., Jurka, J., Kiiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J.R., Murray, J., Nergadze, S.G., Onofrio, R., Pedroni, S., Piras, M.F., Raudsepp, T., Rocchi, M., Røed, K.H., Ryder, O.A., Searle, S., Skow, L., Swinburne, J.E., Syvänen, A.C., Tozaki, T., Valberg, S.J., Vaudin, M., White, J.R., Zody, M.C., Lander, E.S., Lindblad-Toh, K., Broad Institute Genome Sequencing Platform Broad Institute Whole Genome Assembly Team. 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**: 865–867. [[Medline](#)] [[CrossRef](#)]
  30. Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B., Zeggini, E., Lango, H., Lyssenko, V., Timpson, N.J., Burt, N.P., Rayner, N.W., Saxena, R., Ardlie, K., Tobias, J.H., Ness, A.R., Ring, S.M., Palmer, C.N., Morris, A.D., Peltonen, L., Salomaa, V., Davey Smith, G., Groop, L.C., Hattersley, A.T., McCarthy, M.I., Hirschhorn, J.N., Frayling, T.M., Diabetes Genetics Initiative Wellcome Trust Case Control Consortium. 2007. A common variant of HMG2 is associated with adult and childhood height in the general population. *Nat. Genet.* **39**: 1245–1250. [[Medline](#)] [[CrossRef](#)]
  31. Willet, P. 1991. *A History of the General Stud-Book*. Weatherbys, Northants.
  32. Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. 1995. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* **376**: 771–774. [[Medline](#)] [[CrossRef](#)]