—Note—

Evaluating the potential roles of the *Gray* and *Extension* loci in the coat coloration of Thoroughbred racing horses

Takahiro SAKAMOTO[1], Jeffrey A. FAWCETT[2] and Hideki INNAN[2]*

[1]Department of Agriculture, The University of Tokyo, Tokyo 113-8657, Japan
[2]SOKENDAI, The Graduate University for Advance Studies, Kanagawa 240-0193, Japan

Horses have substantial variation in coat color, and the genetic loci responsible for the coat color variations have been well investigated. It has been believed that some color variations should follow a single-locus Mendelian law. Examples include the Gray locus that causes the gray phenotype and the Extension locus that specifies the chestnut phenotype. We reevaluated the roles of the Gray and Extension loci by using a large number of mating records of Thoroughbred racing horses. We showed that the data indeed fits the Mendelian law extremely well for the two loci. Furthermore, we demonstrated that the Extension and Agouti loci might have an additional role in determining the degree of melanin that should distinguish bay, dark bay, and brown.

Key words: *coat color, Mendelian law, Thoroughbred racing horse*

J. Equine Sci.
Vol. 28, No. 2
pp. 61–65, 2017

The domestication of horses resulted in multiple breeds that contain striking phenotypic variations between and within them (see Librado *et al*. [1] for a review). Coat coloration is one of such variations, and we currently observe various colors including chestnut, gray, and black. Even within the popular Thoroughbred racing horses, there is a wide range of variation in coat coloration. For example, the Japan Association for International Racing and Stud Book (JAIRS) recognizes eight colors: chestnut, dark chestnut, gray, white, bay, dark bay, brown, and black. So far, a number of genes have been identified that are involved in the coat coloration system in horses (reviewed by Rieder [3] and Thiruvenkadan *et al.* [6]). For example, Rosengren Pielberg *et al.* [5] found that a duplication of intron 6 of the *STX17* gene causes the gray phenotype, while a missense mutation in the *MC1R* gene is responsible for the chestnut phenotype. These two genes are heavily involved in the coat-coloration system of Thoroughbred racing horses, and it is believed that they follow the simple Mendelian law as illustrated in Fig. 1 [5]. The *Gray* locus (i.e., *STX17*) has two alternative alleles, *G* and *g*, and this locus is epistatic to all other colors except for very rare white horses. *Gray* follows a dominant mode of inheritance; that is, *G*/- creates the gray phenotype and otherwise results in non-gray phenotypes. The *Extension* locus that has two alleles, *E* and *e*, is the next dominant one that distinguishes the chestnut and dark chestnut phenotypes (we refer to the two colors pooled together as chestnut*) from the remaining four colors. It has been demonstrated that the *MC1R* gene, which produces the black pigment melanin, is responsible for the *Extension* locus [2]. *Exten-*

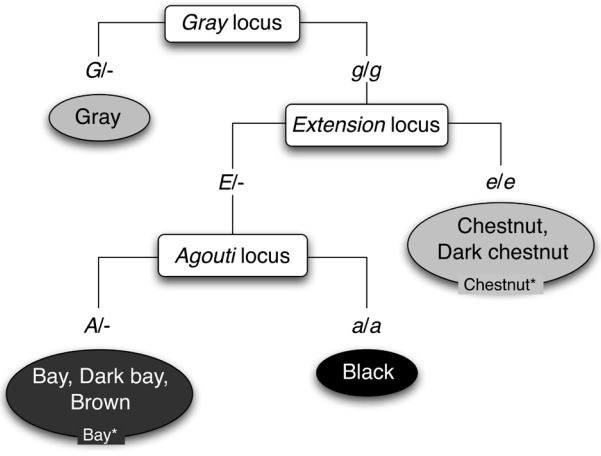

Fig. 1. Illustration of the roles of the *Gray*, *Extension*, and *Agouti* loci in determination of the seven colors (except for white) of the Thoroughbred racing horses.

Received: February 20, 2017
Accepted: May 18, 2017
*Corresponding author. e-mail: innan_hideki@soken.ac.jp

Table 1. The proportion of gray offspring produced

Sire dam	Gray × Gray	Gray × Non-gray	Non-gray × Gray	Non-gray × Non-gray
No. of matings	614	8,482	8,632	136,024
No. of gray foals	454	4,328	4,354	0
Observed proportion	0.739	0.510	0.504	0.000
Expected proportion	0.758	0.508	0.508	0.000

Results for a total of 153,752 matings are shown after excluding 6 cases where either the sire or dam was white.

sion follows a recessive mode of inheritance where *e*/*e* creates the chestnut* phenotype. An additional locus, the *Agouti* locus that has two alleles *A* and *a*, distinguishes the black phenotype from the remaining three colors (bay, dark bay, and brown, which we refer to altogether as bay*). This locus contains the agouti signaling protein (*ASIP*) gene, which controls the amount of melanin produced and also its distribution [4]. This gene has no effect on *e*/*e* chestnut* horses, as there are no black pigments to restrict. On the other hand, *E*/- horses with a functional *A* allele (*A*/-) have a bay* coat color, whereas the amount of melanin produced in horses homozygous for a recessive mutation in this gene (*a*/*a*) cannot be restricted, and these horses therefore have a black coat color. Furthermore, a statistically significant tendency for lighter bay shades in *E/e* horses and darker bay shades in *E/E* horses was found, suggesting that the *Extension* locus may have a role in determining the degree of melanin produced, although this observation was based on a very limited number of horses [4]. The purpose of this note is to explore the potential of the large number of mating records in the Japan Bloodstock Information System (JBIS). To demonstrate its use for genetic analyses, we first reevaluate the roles of the *Gray* and *Extension* loci, for which the data indeed fit the Mendelian law illustrated in Fig. 1 extremely well. Furthermore, we demonstrate that the *Extension* and *Agouti* loci might have an additional role in determining the degree of melanin that should distinguish the different bay* coat colors (bay, dark bay, and brown).

We downloaded data for 153,778 horses that were born between 1995 and 2014 and officially registered in the JBIS (http://www.jbis.or.jp). The data contained the coat color of each horse, the IDs of the sire and dam, and also the coat colors of both parents. We first focused on the *Gray* locus. As the system in Fig. 1 applies only to non-white horses, we excluded 20 white horses, resulting in 153,758 non-white horses. Because 9,137 and 144,621 of them were gray and non-gray horses, respectively, we could roughly estimate the frequency of the *G* allele to be f_G=0.030 from $(1–f_G)^2$=144,621/153,758, assuming the Hardy-Weinberg principle. Given this estimate of f_G=0.030, we obtained the expected proportions for producing the gray phenotype for the four possible mating patterns as summarized in Table 1. The results demonstrate that the observation fits the expectation extremely well. In Fig. 2, the same analysis was applied to the six non-gray phenotypes separately. Again, the observation and expectation are in excellent agreement, indicating that the *Gray* locus is definitely epistatic to the other six color phenotypes. It should also be noted that the 136,024 matings of non-gray × non-gray produced no gray horses, indicating an extremely low de-novo mutation rate from *g* to *G*, and that the registered records (both coat color and pedigree) are very reliable.

Let us next consider the role of the *Extension* locus, where it is predicted that the *e*/*e* homozygote causes the chestnut* phenotype and that *E*/- is responsible for the other four phenotypes (bay* and black). We here used 144,621 non-gray horses, of which 37,709 were chestnut*. Therefore, the frequency of the *E* allele (f_E) was estimated to be 0.489, which was used to compute the expected proportions for producing the chestnut* phenotype for the four possible mating patterns as summarized in Table 2. The results show that the observed proportions are very well explained by the expectations based on f_E. However, when the four different *E*/- phenotypes (bay, dark bay, brown, and black) were analyzed separately, we found significant deviations from the expectations (Fig. 3). The observed proportion of chestnut* is significantly higher than expected for the mating pairs involving the bay phenotype, while it is lower for the mating pairs involving the dark bay and brown phenotypes. The mating pairs involving the black phenotype do not show significant deviations. One possibility is that the association between the *Extension* locus and the chestnut* phenotype is not complete; that is, some *e/e* foals are not chestnut and/or some *E/-* foals are chestnut*. However, this seems unlikely especially considering that all 9,002 matings between chestnut* parents exclusively produced chestnut* foals. Alternatively, this result could also be explained if the number of *E* alleles is associated with the darkness of the bay coat color, as previously suggested [5], in which case the f_E in the different bay* parents should significantly deviate from the expectations. For instance, if the bay parents consist of a larger than expected number with *E/e*

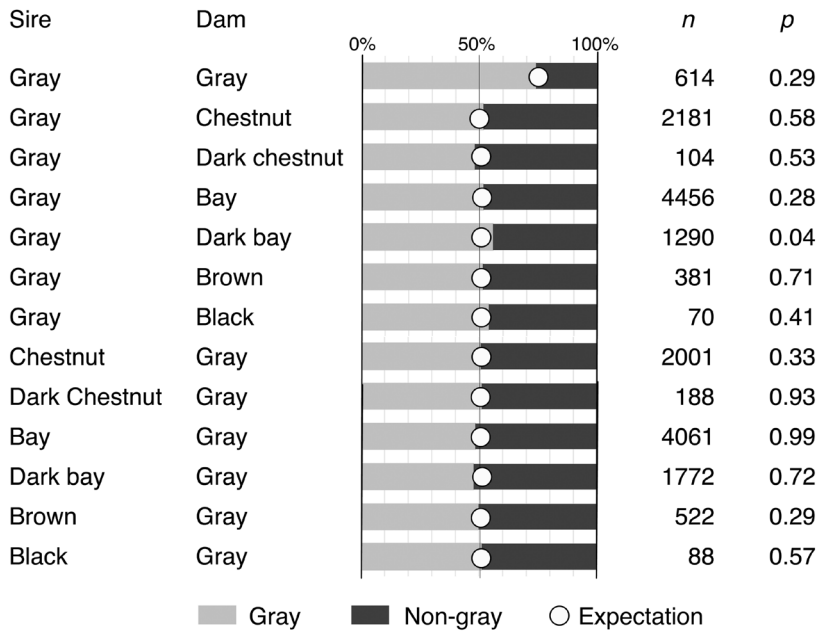

Fig. 2. The proportion of gray offspring produced for 13 mating pairs. *P*-values without correcting for multiple tests are shown.

Table 2. The proportion of chestnut* offspring produced

Sire dam	Chestnut* × Chestnut*	Chestnut* × Bay*/Black	Bay*/Black × Chestnut*	Bay*/Black × Bay*/Black
No. of matings	9,002	25,605	26,172	75,245
No. of chestnut* foals	9,002	8,606	8,998	8,874
Observed proportion	1.000	0.336	0.344	0.118
Expected proportion	1.000	0.338	0.338	0.114

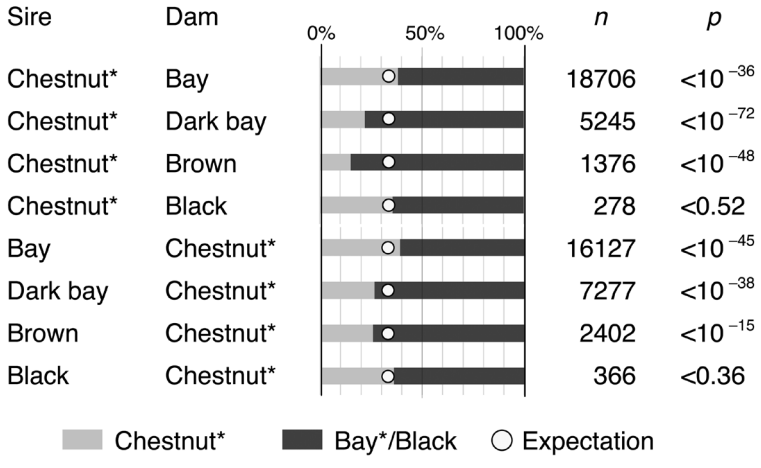

Fig. 3. The proportion of chestnut* offspring produced for 8 mating pairs. *P*-values without correcting for multiple tests are shown.

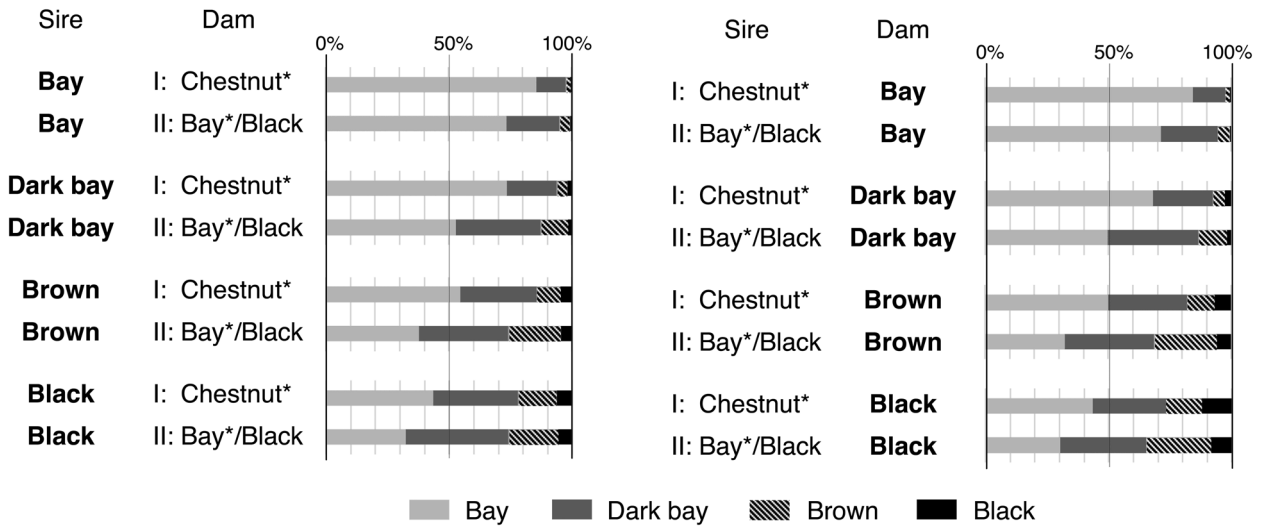

Fig. 4. The relative proportions of the four different bay*/black offspring out of all the bay*/black offspring produced by the patterns I and II (see text for details).

horses compared to *E/E* horses, the proportion of chestnut* foals resulting from matings between bay and chestnut* horses will be larger than the expectations. Likewise, if the dark bay and brown parents consist of a larger than expected number of *E/E* horses compared with *E/e* horses, the proportion of chestnut* foals resulting from the matings between dark bay or brown and chestnut* horses will be smaller than the expectations.

We could not directly compare the genotypes *E*/*E* and *E*/*e,* as we only had phenotype data. Nevertheless, to further test the hypothesis that the *E* locus affects the darkness of the bay* phenotype, we designed an analysis that compares two patterns of mating, I and II, as shown in Fig. 4. To illustrate this analysis, let us use as an example the case where the sire is bay and the dam is chestnut* (*e*/*e*) for pattern I and the sire is bay and the dam is bay*/black. In this comparison, the *E*/- (bay*/black) foals produced by pattern I should all have the *E*/*e* genotype, while the *E*/- (bay*/black) foals produced by pattern II should consist of genotypes *E*/*E* and *E*/*e*. Thus, if *E*/*E* horses are more likely to have darker bay coat colors (dark bay or brown) than *E*/*e* horses, we predict that the proportion of foals with darker bay colors would be larger in pattern II compared with pattern I. The relative proportions of the four *E*/- phenotypes are shown in Fig. 4. The results clearly demonstrate that the proportions of dark bay and brown are increased in pattern II, indicating that *E*/*E* is associated with the increase in the dark bay and brown phenotypes. Fig. 4 shows that this trend is observed for all I vs. II comparisons, consistent with our prediction. Furthermore, we estimated the frequency of *E* (f_E) in each of the four phenotypes, bay, dark bay, brown, and black, to be 0.61, 0.75, 0.77, and 0.64, respectively. Consistent with the results in Fig. 3, the frequency of *E* is larger in the order of brown, dark bay, and bay (excluding black), which correlates with the order of the melanin quantity.

We next considered the role of the *Agouti* locus, where *E/-* horses homozygous for the *a* allele have the black coat color whereas *E/-* horses with a functional *A* allele (*A/-*) have the bay* coat color. The *a* allele is very rare, and we found only three black × black matings. While all of them produced the black phenotype as expected from Fig. 1, the amount of data is too small to have sufficient statistical power to test the Mendelian segregation at the *Agouti* locus. Nevertheless, we can ask whether the number of *A* alleles is associated with the darkness of the bay* coat color using the same approach as Fig. 4. For instance, if the sire is bay (*AA* or *Aa*) and the dam is bay* (*AA* or *Aa*) (pattern I), the bay* foals produced should include both *AA* and *Aa*, whereas if the sire is bay and the dam is black (*aa*) (pattern II), the bay* foals produced should all be *Aa*. Because the *Agouti* locus controls the production of melanin, the *Aa* horses may be more likely to have darker bay colors than *AA* horses. If so, the bay* foals produced by pattern II should contain a larger proportion of darker bay* foals (dark bay or brown) compared with those produced by pattern I. Indeed, as shown in Fig. 5, the proportion of lighter bay foals (bay) is much larger in pattern I, whereas the proportion of darker bay foals is much larger in pattern II for all six comparisons. Our estimates of the frequency of *A* (f_A) in the three bay* phenotypes are consistent with the prediction (f_A=0.95, 0.82, and 0.58 for bay, dark bay, and brown, respectively). Although the sample size is a lot smaller than in Fig. 4, this

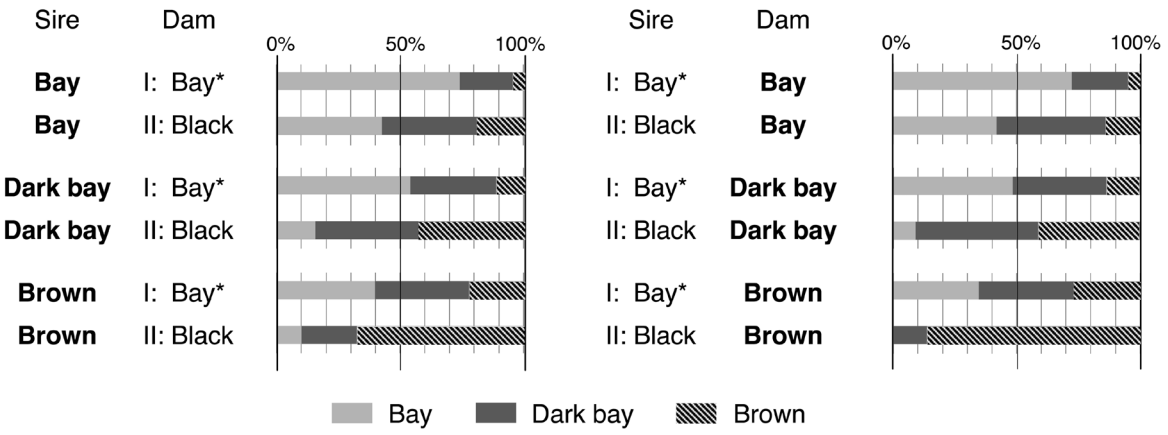

Fig. 5. The relative proportions of the three different bay* offspring produced by the patterns I and II (see text for details).

result indicates that the offspring with the *A* allele are likely to have a lower melanin quantity.

In summary, we reevaluated the potential roles of two well-studied loci, *Gray* and *Extension*, by taking advantage of a very large amount of data in the JBIS. For the *Gray* locus, the observed data were in perfect agreement with the prediction of the simple Mendelian law. The *Extension* locus also explained the segregation of chestnut* and bay*/black. Furthermore, we demonstrated that the number of *E* alleles present should have some effect on the segregation of the three quantitative bay* phenotypes (bay, dark bay, and brown), perhaps by affecting the level of melanin produced. There could be at least two possibilities to explain this result. First, the product of the *MC1R* gene might have a direct effect on the amount of melanin produced. Second, somewhere very close to the *MC1R* gene, there might be another gene involved in the control of melanin. A similar role of the *Agouti* locus was also indicated, suggesting that the pattern of coat coloration involves a complex system with a number of loci. This work thus demonstrates the power of large data to reveal as-yet unknown roles of coat color genes, and also potentially genes behind other phenotypes.

References

1. Librado, P., Fages, A., Gaunitz, C., Leonardi, M., Wagner, S., Khan, N., Hanghøj, K., Alquraishi, S.A., Alfarhan, A.H., Al-Rasheid, K.A., Der Sarkissian, C., Schubert, M., and Orlando, L. 2016. The evolutionary origin and genetic makeup of domestic horses. *Genetics* **204**: 423–434. [Medline] [CrossRef]
2. Marklund, L., Moller, M.J., Sandberg, K., and Andersson, L. 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. *Mamm. Genome* **7**: 895–899. [Medline] [CrossRef]
3. Rieder, S. 2009. Molecular tests for coat colours in horses. *J. Anim. Breed. Genet.* **126**: 415–424. [Medline] [CrossRef]
4. Rieder, S., Taourit, S., Mariat, D., Langlois, B., and Guérin, G. 2001. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mamm. Genome* **12**: 450–455. [Medline] [CrossRef]
5. Rosengren Pielberg, G., Golovko, A., Sundström, E., Curik, I., Lennartsson, J., Seltenhammer, M.H., Druml, T., Binns, M., Fitzsimmons, C., Lindgren, G., Sandberg, K., Baumung, R., Vetterlein, M., Strömberg, S., Grabherr, M., Wade, C., Lindblad-Toh, K., Pontén, F., Heldin, C.H., Sölkner, J., and Andersson, L. 2008. A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nat. Genet.* **40**: 1004–1009. [Medline] [CrossRef]
6. Thiruvenkadan, A. K., Kandasamy, N., and Panneerselvam, S. 2008. Coat colour inheritance in horses. *Livest. Sci.* **117**: 109–129. [CrossRef]