Should we use the SNP linked to DMRT3 in genomic evaluation of French trotter?

Sophie Brard, Anne Ricard

JAS9224

Should we use the single nucleotide polymorphism linked to *DMRT3* in genomic evaluation of French trotter?[1]

S. Brard*†‡[2] and A. Ricard§#

*INRA, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), 31326 Castanet-Tolosan, France; †Université de Toulouse, INP, ENSAT, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), 31326 Castanet-Tolosan, France; ‡Université de Toulouse, INP, ENVT, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), 31076 Toulouse, France; §INRA, UMR 1313, 78352 Jouy-en-Josas, France; and #IFCE, Recherche et Innovation, 61310 Exmes, France

ABSTRACT: An A/C mutation responsible for the ability to pace in horses was recently discovered in the *DMRT3* gene. It has also been proven that allele C has a negative effect on trotters' performances. However, in French trotters (FT), the frequency of allele A is only 77% due to an unexpected positive effect of allele C in late-career FT performances. Here we set out to ascertain whether the genotype at SNP *BIEC2-620109* (linked to *DMRT3*) should be used to compute EBV for FT. We used the genotypes of 630 horses, with 41,711 SNP retained. The pedigree comprised 5,699 horses. Qualification status (trotters need to complete a 2,000-m race within a limited time to begin their career) and earnings at different ages were precorrected for fixed effects and evaluated with a multitrait model. Estimated breeding values were computed with and without the genotype at SNP *BIEC2-620109* as a fixed effect in the model. The analyses were performed using pedigree only via BLUP and using the genotypes via genomic BLUP (GBLUP). The genotype at SNP *BIEC2-620109* was removed from the file of genotypes when already taken into account as a fixed effect. Alternatively, 3 groups of 100 candidates were used for validation. Validations were also performed on 50 random-clustered groups of 126 candidates and compared against the results of the 3 disjoint sets. For performances on which *DMRT3* has a minor effect, the coefficients of correlation were not improved when the genotype at SNP *BIEC2-620109* was a fixed effect in the model (earnings at 3 and 4 yr). However, for traits proven strongly related to *DMRT3*, the accuracy of evaluation was improved, increasing +0.17 for earnings at 2 yr, +0.04 for earnings at 5 yr and older, and +0.09 for qualification status (with the GBLUP method). For all traits, the bias was reduced when the SNP linked to *DMRT3* was a fixed effect in the model. This work finds a clear rationale for using the genotype at *DMRT3* for this multitrait evaluation. Genomic selection seemed to achieve better results than classic selection.

Key words: *DMRT3*, genomic selection, horse, major gene, single nucleotide polymorphism, trotter

J. Anim. Sci. 2015.93:1–9
doi:10.2527/jas2015-9224

INTRODUCTION

Andersson et al. (2012) discovered a major gene affecting locomotion in horses. A stop mutation in *DMRT3* is strongly associated with ambling gaits, which are very comfortable gaits that some breeds naturally have or are easily able to learn due to a genetic predisposition, in addition to the usual gaits (walk, trot, and canter). This mutation is caused by a single base change: the wild-type allele C is replaced by the mutant allele A. Promerová et al. (2014) found that the mutated allele was also fixed in many breeds dedicated to trot races but missing in breeds selected for gallop races. A feature of trot races is that horses that break stride are disqualified, and so trotters have been selected on their ability to trot easily at high speed. The mutated allele was proven to have a positive effect on racing performances in Swedish standardbred trotters and is fixed in American standardbred trotters. Nevertheless, in French trotters

[1]This work was funded by the SelGen metaprogramme and by the French Institute for Horses and Riding (Institut Français du Cheval et de l'Equitation, IFCE).

[2]Corresponding author: sophie.brard@toulouse.inra.fr

Received April 21, 2015.
Accepted July 27, 2015.

(**FT**), the frequency of the mutation is only 77%. Ricard (2015) studied the effect of the genotype at SNP *BIEC2-620109* (linked to *DMRT3*; C-C and A-T**[AU: please confirm changes]**) and confirmed positive effects of the mutation on ability to trot easily and on earnings through most of the career of FT. Nevertheless, the greater earnings are obtained in prestigious events that are mainly raced at 5 yr and older by only few horses. The wild-type allele in heterozygotes had a positive and highly significant effect on these late earnings ($P < 0.001$), which justified its frequency in a long-term-selected breed such as FT. Our objective is to ascertain whether using the genotype at this same SNP in high linkage disequilibrium (**LD**) with *DMRT3* would enable a better estimation of breeding values for performances by FT in harness races. We also considered the effect of the method used to obtain the relationship matrix: expected relationships based on pedigree or realized relationships based on genotypes. This work, therefore, reports the results of a first attempt at genomic selection in FT.

MATERIALS AND METHODS

Phenotypes

The races studied in this work are harness races, in which the horse pulls an ultralight roadster called a sulky. Performances in these races may be analyzed through different traits. The first step in the career of a trotter is to pass a qualification test to gain the right to compete in races. The qualification test consists of a 2,000-m race that has to be completed within a limited time allocation, which can change every year depending on improvement of the racing performances of the whole population and which is also dependent on the age of the horse. About 40% of a given generation passes the test. Qualification is, therefore, an important trait for 2 reasons: first, it is a relatively highly heritable (Table 1) and early trait, and second, it means that horses that will race are a selected sample of their generation. The subsequent career of a trotter can then be considered at 3 stages. The first stage is racing as a 2 yr old. This is an early stage trait as only 20% of the horses racing at 3 yr started at 2 yr. The second stage, racing as a 3 or 4 yr old, is the crux of a trotter's career. Few of them will go on to make the third stage, that is, racing at 5 yr and older: one-third of the horses stop racing after 4 yr old and another one-third will stop each year that follows. Horses win money prizes depending on their ranking. Most of the time, only the first 7 horses receive a prize. The first horse earns half of the total prize, and then the second receives half of the remaining money, and so on down to the seventh horse. The next 9 horses are ranked but do not earn money. According to Thiruvenkadan et al. (2009), performances in trot racing can be studied using the logarithm of annual earnings divided by the annual number of finished races (**LnE**), assuming a horse can be disqualified if it breaks stride. Here, we study LnE based on these 3 stages: early earnings at 2 yr, peak of career at 3 yr and 4 yr separately, and late earnings, which will include all prizes at 5 yr and older. Heritability of earnings is moderate (around 0.30; Table 1). The records for all these traits were provided by the Society for the Promotion of French Horse Breeding (Société d'encouragement à l'élevage du cheval français; **[AU: please provide city]**, France) and by the French Institute for Horses and Riding (Institut français du cheval et de l'équitation; **[AU: please provide city]**, France). Although records were available for all horses that took part in French races from 1996 to 2011, the data were truncated: records were kept only for horses born between 1994 and 2008 to have a sufficient amount of information on all horses.

Table 1. Heritability (diagonal), genetic correlation (upper triangle), and residual correlation (lower triangle) for logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages and qualification status as per Ricard (2015)

	LnE at				Qualification
Trait	2 yr	3 yr	4 yr	≥5 yr	status
LnE at 2 yr	0.28	0.85	0.76	0.56	0.48
LnE at 3 yr	0.29	0.32	0.91	0.81	0.61
LnE at 4 yr	0.12	0.27	0.25	0.92	0.47
LnE at 5 yr and older	0.14	0.23	0.41	0.26	0.44
Qualification status	0.00[1]	0.00[1]	0.00[1]	0.00[1]	0.56

[1]Residual correlations between qualification status and LnE were fixed to 0.

Horses

The blood samples of 623 horses had been previously collected for a genomewide association study on osteochondrosis in FT (Teyssèdre et al., 2012). The horses were recruited between 2008 and 2010 at French Center for Imaging and Research on Equine Locomotor Disorders (**[AU: please provide city]**, France) and at a few veterinary clinics. These horses are not exactly a random sample of the population, and they have globally better performances than other trotters (79% of them were qualified whereas only about 40% of a generation usually makes the cut). Another 59 horses were genotyped, giving a total of 682 genotyped horses available. Finally, 630 were retained, based on the availability of their records for racing performances: they were born between 1994 and 2008 and their performances were recorded between 1996 and 2011. Of the retained sample, 41% were females and 61 horses were sires. Looking in from an alternative perspective, 300 genotyped horses

had their sire genotyped. As sires are selected on their own performances, their presence in the sample also explains the better performances recorded. The 630 trotters were included in a pedigree of 5,699 horses.

Genotypes

The genotypes were obtained using the Illumina Equine SNP50 BeadChip (**[AU: please provide the full name of manufacturer and manufacturer's city and state (or city and country if outside the U.S.)]**). A quality control of SNP genotypes based on minimum allele frequency (≥1%), genotype assessment rate (≥80%), and Hardy–Weinberg equilibrium (*P*-value for the test $>10 \times 10^{-8}$) was performed, and 41,711 SNP were retained. The chip marker that had the strongest LD with the mutation identified in *DMRT3* (Andersson et al., 2012) was the SNP *BIEC2-620109*. Promerová et al. (2014) estimated a LD of 0.91 between this SNP and the *DMRT3* mutation in a population of 2,749 horses including 59 FT. The C allele at SNP *BIEC2-620109* is associated with the C allele in the mutation, whereas the T allele of the SNP is associated with the A allele. The frequencies for the 3 genotypes among the retained horses were 56% for TT, 39% for TC, and 5% for CC.

Statistical Models

Qualification status and LnE at different ages were studied by Ricard (2015) in the same population with the objective of assessing the effect of SNP *BIEC2-620109* on performances in trot racing. Ricard (2015) used the following model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e},$$

in which **y** is the performance vector, **b** is the fixed effect vector that combines gender and year of birth, **a** is the vector of random polygenic values, and **e** is the vector of residuals. $V(\mathbf{a}) = \mathbf{A}\sigma^2_a$, in which **A** is the relationship matrix and σ^2_a is the **[AU: please define σ²a]** and $V(\mathbf{e}) = \mathbf{I}\sigma^2_e$, in which σ^2_e is the **[AU: please define σ²e]**. **X** and **Z** are incidence matrices. Heritability, genetic correlations, and residual correlations between traits had been estimated in a multitrait model using more than 173,000 FT, with about 64,000 of them qualified. Here we used performances precorrected for fixed effects according to the estimations obtained with this model. This same approach has already been used by Pribyl et al. (2010). We realized a multitrait estimation of breeding values for qualification status and LnE at different ages to exploit the genetic correlations between traits (Table 1). Note that qualification status was first a binary variable (0 = unqualified and 1 = qualified), but the correction for fixed effects turned it into a continuous trait. Qualification status is important in the multitrait evaluation because it allows the use of horses that are not qualified and have not yet posted earnings. It has been demonstrated that it is important to use those horses without earnings in the estimation of breeding values to reduce the bias (Klemetsdal, 1992; Árnason, 1999).

Our first objective was to assess whether the genotype of the SNP linked to *DMRT3* should be used to compute the EBV, as it has been shown that QTL are a useful source of information for animal selection (Soller and Beckmann, 1983). Our second objective was to check whether genomic selection should be preferred over pedigree-based selection. For comparison of classic vs. genomic selection, 2 methods were used to compute the EBV: a BLUP and a genomic BLUP (**GBLUP**). On the one hand, EBV were calculated using the pedigree, and on the other hand, genomic EBV were calculated using the relationship matrix deduced from the horse genotypes. The corresponding statistical model is

$$\mathbf{y^*} = \mathbf{1}\mu + \mathbf{Zg} + \boldsymbol{\varepsilon},$$

in which $\mathbf{y^*}$ is the vector of precorrected performances of the 630 horses, **1** is a vector of ones, μ is the overall mean, **g** is a random vector of additive genetic values, and $\boldsymbol{\varepsilon}$ is a vector of residuals. In BLUP, $\text{var}(\mathbf{g}) = \mathbf{A}\sigma^2_g$, in which **A** is the pedigree-based relationship matrix and σ^2_g is the additive genetic variance. In GBLUP, $\text{var}(\mathbf{g}) = \mathbf{G}\sigma^2_g$ and **G** is the genomic relationship matrix as defined by VanRaden (2008). **Z** is an incidence matrix. This model allows the comparison of pedigree-based and marker-based evaluations. To test the value of using the SNP linked to *DMRT3* for the estimation of breeding values, we modified the model by adding a fixed effect. The model, therefore, became

$$\mathbf{y^*} = \mathbf{1}\mu + \mathbf{Zg} + \mathbf{W}\beta + \boldsymbol{\varepsilon},$$

in which β is the vector of the fixed effect of the 3 genotypes and **W** is the incidence matrix for the horse's genotype at this SNP. When this model was coupled to the GBLUP method, the SNP *BIEC2-620109* was removed from the file of genotypes that was used to compute **G**; therefore, the genotype at this SNP was used only once in the model. Estimated breeding values were computed using BLUPF90 (Misztal et al., 2002).

Validation

Comparison between methods was based on cross-validation. The file was divided into 2 populations: reference population and candidate population. Estimated breeding values were calculated including performances

Table 2. Correlation coefficients between EBV and performances for the corresponding traits: logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages or qualification status

Trait[1]	Effective number of candidates[2]	The SNP linked to *DMRT3* is not used		The SNP linked to *DMRT3* is a fixed effect	
		BLUP	GBLUP[3]	BLUP	GBLUP
LnE at 2 yr	38	0.19	0.27	0.40	0.44
LnE at 3 yr	171	0.21	0.29	0.19	0.27
LnE at 4 yr	134	0.28	0.34	0.22	0.26
LnE at 5 yr and older	83	0.43	0.41	0.44	0.45
Qualification status	300	0.13	0.16	0.27	0.25

[1]Estimated breeding values were computed in a multitrait analysis, with 4 combinations of models and methods, and for 3 nonoverlapping validation data sets of 100 candidates each.

[2]Results are presented for all candidates pooled together. Qualification status was the only trait for which all candidates had a performance record. For LnE, the nonqualified horses had missing values and could not, therefore, be candidates.

[3]GBLUP = genomic BLUP.

of the reference population only, and the validation criteria were then based on relationships between EBV and performances of the candidate population. For the accuracy, we used the correlation coefficient between the candidates' EBV and their realized performances. For the bias, we used the regression of the realized performances of candidates on their EBV. When the model included the SNP *BIEC2-620109*, the total EBV of a horse was the solution for the animal effect summed to the solution for its genotype at SNP *BIEC2-620109*.

The candidates had to meet the following requirements:

- have at least one record (not necessarily a record for each of the traits studied in the multitrait analysis),
- be genotyped,
- be the son of a genotyped sire, itself having at least one record, and
- not have any progeny with records.

Because the performances were precorrected for fixed effects including year of birth, this information was not used to design the reference and validation samples.

A total of 300 horses were potential candidates. If EBV had been computed simultaneously for all of them, the reference population would have been reduced to 330 individuals. To use enough information to compute candidates' EBV, a 3-fold cross-validation was used. Candidates were randomly divided into 3 nonoverlapping training data sets of equal size (100 horses), and each group of candidates was then evaluated one by one, with the 2 other groups included in the reference population. This method has already been used in dairy cattle (Luan et al., 2009), beef cattle, (Saatchi et al., 2011), and pine (Resende et al., 2012). The advantage of this method was to guarantee that EBV of the trotters in the validation sample would be computed using a reference population counting at least 5 times more individuals, in line with recommended practice (Legarra et al., 2008). Accuracy and bias were then computed for the 3 groups of candidates together.

To quantify the SE due to sampling of the results of the 3-fold cross-validation, EBV were also computed for 50 random groups of 126 candidates, yielding four-fifths genotyped horses with performances in the reference population and one-fifth in the validation population. In this case, the groups were obviously overlapping, and candidates had several EBV as they could be picked from several data sets for validation. We therefore computed the accuracy and the bias separately for each of the 50 validation data sets. The accuracy and bias obtained with the 3-fold cross-validation were compared against the distributions of the correlation and regression coefficients obtained for the 50 overlapping data sets.

RESULTS

Validation on 3 Nonoverlapping Data Sets

The accuracy and the bias computed for the 300 candidates evaluated in the 3-fold cross-validation are shown in Tables 2 and 3. All the horses of the validation population had a record for qualification status but may have had earnings for only some of the years of their career or no earning at all. Therefore, the correlation and regression coefficients were computed for different numbers of candidates depending on the number of trotters that truly had a record for the trait. Early and late earnings (LnE at 2 yr and LnE at 5 yr and older) were the traits that had lower numbers of candidates, at 38 and 83, respectively. Logarithm of annual earnings divided by the LnE at 3 and 4 yr had higher numbers of candidates (at 134 and 171, respectively).

Early and late earnings and qualification status achieved greater accuracies when the SNP *BIEC2-620109* was included as a fixed effect in the model (Table 2). The superiority of this model was more obvious for LnE at 2 yr (+0.21 for BLUP and +0.17 for GBLUP) and qualification status (+0.14 for BLUP and +0.09 for GBLUP) than for LnE at 5 yr and older (+0.01

Table 3. Regression coefficients of the performances on EBV for the corresponding traits: logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages or qualification status

Trait	Effective number of candidates[1]	The SNP linked to *DMRT3* is not used		The SNP linked to *DMRT3* is a fixed effect	
		BLUP	GBLUP[3]	BLUP	GBLUP
LnE at 2 yr	38	1.70	1.55	1.39	1.44
LnE at 3 yr	171	1.28	1.33	1.02	1.16
LnE at 4 yr	134	1.67	1.65	1.04	1.06
LnE at 5 yr and older	83	2.29	1.83	1.35	1.25
Qualification status	300	0.74	0.60	1.05	0.78

[1]Estimated breeding values were computed in a multitrait analysis, with 4 combinations of models and methods, and for 3 nonoverlapping validation data sets of 100 candidates each.

[2]Results are presented for all 300 candidates pooled together. Qualification status was the only trait for which all candidates had a performance record. For LnE, the nonqualified horses had missing values and could not, therefore, be candidates.

[3]GBLUP = genomic BLUP.

for BLUP and +0.04 for GBLUP). This result was consistent with the significant effect of the SNP linked to *DMRT3* on those traits as originally evidenced by Ricard (2015). With this model, GBLUP provided slightly greater accuracies for LnE at 2 yr (+0.04) and 5 yr and older (+0.01), whereas for the qualification status, BLUP gave marginally better results than GBLUP (+0.02).

For LnE at 3 and 4 yr, greater accuracies were achieved when the SNP *BIEC2-620109* was not a fixed effect of the model, particularly for LnE at 4 yr (+0.12). This is consistent with previous results of Ricard (2015): the SNP linked to *DMRT3* is thought to have a less significant effect on these traits, so the accuracy is not improved when the SNP *BIEC2-620109* is added in the model. For these 2 traits, the superiority of using GBLUP over BLUP was ascertained (+0.08 for accuracy for LnE at 3 yr and +0.06 at 4 yr).

For traits that achieved better accuracy when the SNP linked to *DMRT3* was in the model (early and late earnings and qualification status), the regression coefficients closest to 1 were also obtained with this model (Table 3). For LnE at 3 and 4 yr, the regression coefficients were nearly unbiased when the SNP linked to *DMRT3* was in the model, although their coefficient of correlation was not improved.

Validation on 50 Overlapping Groups of Candidates

Figure 1 shows the distribution of the accuracy achieved by the 50 random-clustered validation data sets. Like for the nonoverlapping groups, the number of effective candidates for each trait in one group was different depending on availability of records. The numbers of effective candidates are shown in Table 4. The distributions of accuracy for LnE at 2 yr and qualification status clearly showed the superiority of the model including the SNP linked to *DMRT3* as a fixed effect. For LnE at 5 yr and older, the superiority of this model was less patent, although the distributions for the model including the SNP *BIEC2-620109* visibly achieved slightly better accuracies. For LnE at 3 yr, the distributions of accuracies for both models were very similar, although the GBLUP approach achieved greater accuracies than the BLUP method. For LnE at 4 yr, the shapes of the distributions of accuracy did not single out a better model or a better method. Note in Fig. 1 that the average values of accuracies among the 50 validation data sets were very close to the accuracies computed on the 3 nonoverlapping validation data sets.

Figure 2 charts the distributions of bias. For all traits, the distributions were closest to 1 when the model included the SNP linked to *DMRT3* as a fixed effect. This result was consistent with observations on the validation based on the 3 nonoverlapping groups. For these models including SNP *BIEC2-620109*, GBLUP seemed to more often achieve unbiased estimations than BLUP for LnE at 2 yr, LnE at 5 yr and older, and qualification status. The BLUP method seemed to achieve regression coefficients closer to 1 for LnE at 3 yr. For LnE at 4 yr, neither BLUP nor GBLUP emerged as superior in terms of the bias. Once again, these results were consistent with the difference between methods evidenced with the 3 nonoverlapping groups.

Table 4. Effective number of candidates in the 50 random-clustered validation data sets for the logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages and qualification status

Trait	Effective number of candidates		
	Mean	Minimum	Maximum
LnE at 2 yr	16	9	21
LnE at 3 yr	72	64	78
LnE at 4 yr	56	47	64
LnE at 5 yr and older	35	28	43
Qualification status	126	126	126

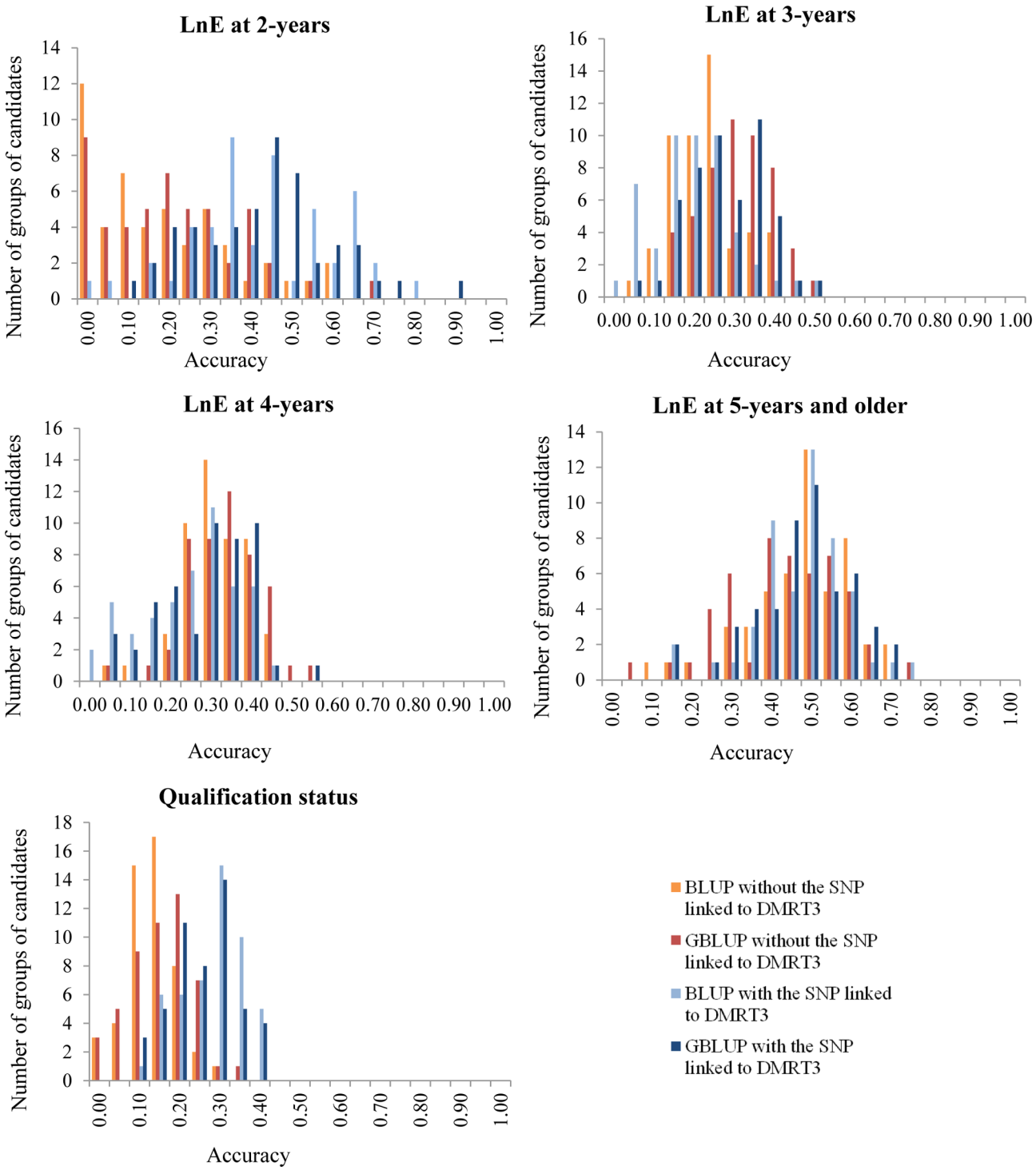

Figure 1. Distributions of correlation coefficients between the EBV and the corresponding performances for 50 randomly clustered validation data sets of 126 candidates each. Estimated breeding values were computed in a multitrait evaluation including logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages and qualification status, with 4 combinations of models and methods. GBLUP = genomic BLUP. **[AU: In the figure, please change "-years" to " yr" (yr with a space between the numeral and the abbreviation yr). Please make DMRT3 italic.]**

DISCUSSION

The results of our study were consistent with the results of Ricard (2015). On one hand, accuracy was improved for LnE at 2 yr, LnE at 5 yr and older, and qualification status when the genotype at SNP *BIEC2-620109* was included in the model. Ricard (2015) found that the genotype with the SNP linked to *DMRT3* had a very highly significantly different effect ($P < 0.001$) on these 3 traits. In our estimation of genotype effects with the GBLUP method based on the 50 validation groups, the difference of effect of genotype

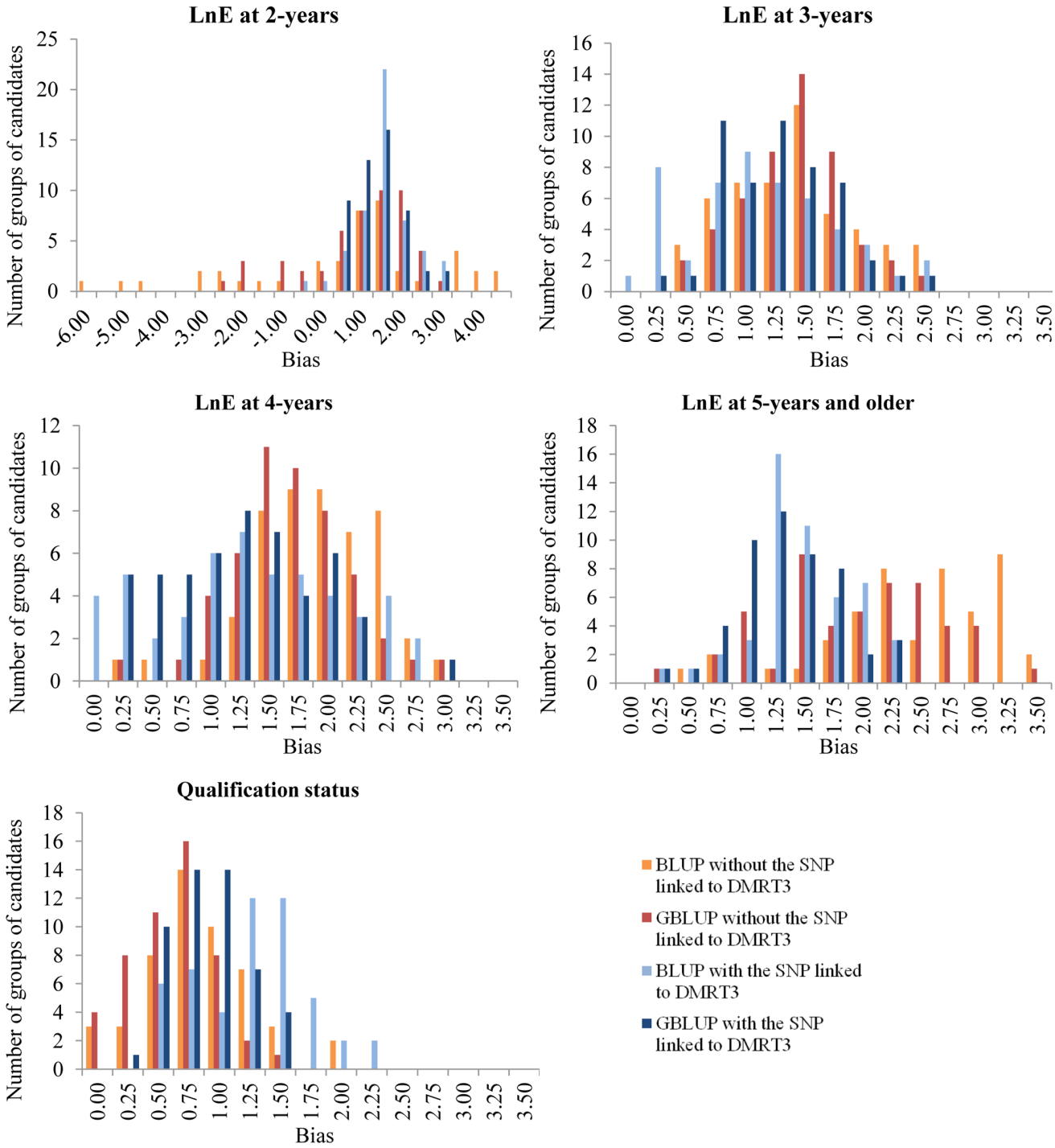

Figure 2. Distributions of regression coefficients of the performances on EBV for the corresponding traits for 50 randomly clustered validation data sets of 126 candidates each. Estimated breeding values were computed in a multitrait evaluation including logarithm of annual earnings divided by the annual number of finished races (LnE) at different ages and qualification status, with 4 combinations of models and methods. GBLUP = genomic BLUP. **[AU: In the figure, please change "-years" to " yr" (yr with a space between the numeral and the abbreviation yr). Please make DMRT3 italic.]**

CT compared with TT was very highly significant for LnE at 2 yr ($P < 0.001$, –0.84 phenotypic SD) and highly significant for LnE at 5 yr and older ($P < 0.01$, +0.44 phenotypic SD) and the difference of effect of genotype CC compared with TT was highly significant for qualification status ($P < 0.01$, –0.70 phenotypic SD). On the other hand, the accuracy for LnE at 3 and 4 yr was slightly lower when the SNP linked to *DMRT3* was included in the model. This could be due to a weak effect of SNP *BIEC2-620109* on these traits: no useful information is added for the computation of the corresponding EBV.

For qualification status, the accuracy reached with the GBLUP method coupled with the model with the

SNP linked to *DMRT3* may be considered low (0.25), as this trait has the greatest h^2 (0.56) and the greatest number of horses with recorded performances. Qualification status is a discrete trait, so we had to use averaged performances for the multitrait evaluation, and the correlation coefficient that we used might not be the most suitable way of measuring accuracy for this kind of phenotype. However, it nevertheless made it possible to observe an increase in accuracy due to the addition of the effect of SNP *BIEC2-620109* in the model, even if we can suppose that accuracies for this trait are generally underestimated with the coefficient of correlation used.

Logarithm of annual earnings divided by the LnE at 5 yr and older obtained quite high accuracies in models both with and without the SNP linked to *DMRT3*. This result was unexpected according to h^2 and the number of horses that do have performances for this trait. These high values may be due to selection.

Even if adding qualification status in the multitrait evaluation is supposed to reduce bias, high regression coefficients were obtained for LnE at 2 and 5 yr and older. This may be due to relative overselection of those horses that do have performances for these traits.

The distributions obtained for accuracy and bias with 50 randomly selected validation data sets showed that very different values could be obtained for both accuracy and bias depending on the group of candidates, which illustrates the SE of accuracy and bias of our cross-validation. A genetic evaluation associating genomic EBV and the effect of a major gene has already been performed in dairy cattle by Hayr et al. (2013) via a single-trait evaluation, whereas we worked on a multitrait evaluation. Hayr et al. (2013) computed EBV for fat yield with the genotype at *DGAT1* in the model and found that the better results were obtained when the major gene was considered as a fixed effect, with no improvement of accuracy when the genotype of the major gene was imputed for all animals. Zhang et al. (2010) developed a different method: they realized a weighted genomic evaluation with a trait-specific marker-derived relationship matrix and achieved better accuracy for traits depending on a major gene compared with BLUP and GBLUP. In their method, the marker-derived relationship matrix is different from the realized relationship matrix used in GBLUP, because a greater weight is attributed to loci depending on the genetic variance they explain. Zhang et al. (2010) did not compare their weighted GBLUP to the method used here (a GBLUP with the major gene as a fixed effect in the model), but there is every reason to believe that this method could be very valuable in trotters given how *DMRT3* has a strong effect on performances. Nevertheless, as discussed earlier, the evaluation of breeding values of trotters is a multiple-trait model that should include qualification status in addition to LnE, and it considers the earnings obtained in each year of a horse's career. As the effect of SNP *BIEC2-620109* is different for each of the traits that we evaluated simultaneously, this would imply deriving a weighted relationship matrix for each of the traits and implementing **[AU: please confirm chang]** a method able to take into account these different matrices, likely resulting in longer computation times.

For now, we recommend using a model with the SNP linked to *DMRT3* as a fixed effect. This model achieved better accuracies for LnE at 2 yr, LnE at 5 yr and older, and qualification status. The drop in accuracy when the SNP linked to *DMRT3* was added in the model was low for LnE at 3 yr but higher for LnE at 4 yr. However, the model including the SNP linked to *DMRT3* as a fixed effect had the advantage of being less biased than the model without the SNP. With the model including the SNP linked to *DMRT3*, we recommend using GBLUP, as it yields greater accuracies than BLUP for all traits except qualification status, which was slightly more accurate with BLUP. A combination of a GBLUP approach with a model including the effect of the major gene, therefore, looks like a good compromise for estimating breeding values in FT. The way to use these breeding values remains open to discussion. Synthetic indexes may be produced with different weights for each trait depending on the breeder's objectives. It would be possible to select horses with a large weight on qualification status and LnE at 3 and 4 yr old to produce horses that would be easy to qualify. We could also imagine an index with a higher weight on performances at 5 yr and older to select horses that would be harder to qualify but expected to perform better later in their career in prestigious high-pay-off races. This type of selection would entail breeding for the C allele. As the CC genotype has a negative effect on all traits, a strategy should be defined to keep this allele in heterozygous horses while minimizing the frequency of the homozygotes in the population.

LITERATURE CITED

Andersson, S. L., M. Larhammar, F. Memic, H. Wootz, D. Schwochow, C. J. Rubin, K. Patra, T. Arnason, L. Wellbring, G. Hjälm, F. Imsland, J. L. Petersen, M. E. McCue, J. R. Mickelson, G. Cothran, N. Ahituv, L. Roepstorff, S. Mikko, A. Vallstedt, G. Lindgren, L. Andersson, and K. Kullander. 2012. Mutations in *DMRT3* affect locomotion in horses and spinal circuit function in mice. Nature. 488:642–646. doi:10.1038/nature11399.

Árnason, T. 1999. Genetic evaluation of Swedish standard-bred trotters for racing performance traits and racing status. J. Anim. Breed. Genet. 116:387–398. doi:10.1046/j.1439-0388.1999.00202.x.

Hayr, M. K., M. Saatchi, D. L. Johnson, and D. J. Garrick. 2013. Increasing the accuracy of genomic predictions of fat yield in New Zealand Holstein Friesians using *DGAT1* genotypes. J. Dairy Sci. 96(Suppl. 1):618–619 (Abstr.).

Klemetsdal, G. 1992. Estimation of genetic trend in racehorse breeding. Acta Agric. Scand., Sect. A 42:226–231. doi:10.1080/09064709209410133.

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

Luan, T., J. A. Woolliams, S. Lien, M. Kent, M. Svendsen, and T. H. E. Meuwissen. 2009. The accuracy of genomic selection in Norwegian Red cattle assessed by cross-validation. Genetics. 183:1119–1126. doi:10.1534/genetics.109.107391.

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). In: **[AU: please provide any editors]** Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France. p. 1–2.

Pribyl, J., V. Rehout, J. Citek, and J. Pribylova. 2010. Genetic evaluation of dairy cattle using a simple heritable genetic ground. J. Sci. Food Agric. 90:1765–1773. doi:10.1002/jsfa.4041.

Promerová, M., L. S. Andersson, R. Juras, M. C. T. Penedo, M. Reissmann, T. Tozaki, R. Bellone, S. Dunner, P. Hořín, F. Imsland, P. Imsland, S. Mikko, D. Modrý, K. H. Roed, D. Schwochow, J. L. Vega-Pla, H. Mehrabani-Yeganeh, N. Yousefi-Mashouf, E. G. Cothran, G. Lindgren, and L. Andersson. 2014. Worldwide frequency distribution of the '*Gait keeper*' mutation in the *DMRT3* gene. Anim. Genet. 45:274–282. doi:10.1111/age.12120.

Resende, M. F. R., Jr., P. Muñoz, M. D. V. Resende, D. J. Garrick, R. L. Fernando, J. M. Davis, E. J. Jokela, T. A. Martin, G. F. Peter, and M. Kirst. 2012. Accuracy of genomic selection methods in a standard data set of Loblolly Pine (*Pinus taeda* L.). Genetics. 190:1503–1510. doi:10.1534/genetics.111.137026.

Ricard, A. 2015. Does heterozygosity at the *DMRT3* gene make French trotters better racers? Genet. Sel. Evol. 47:10. doi:10.1186/s12711-015-0095-7.

Saatchi, M., M. C. McClure, S. D. McKay, M. M. Rolf, J. W. Kim, J. E. Decker, T. M. Taxis, R. H. Chapple, H. R. Ramey, S. L. Northcutt, S. Bauck, B. Woodward, J. C. M. Dekkers, R. L. Fernando, R. D. Schnabel, D. J. Garrick, and J. F. Taylor. 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genet. Sel. Evol. 43:40. doi:10.1186/1297-9686-43-40.

Soller, M., and J. S. Beckmann. 1983. Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. 67:25–33. doi:10.1007/BF00303917.

Teyssèdre, S., M. C. Dupuis, G. Guérin, L. Schibler, J. M. Denoix, J. M. Elsen, and A. Ricard. 2012. Genome-wide association studies for osteochondrosis in French Trotter horses. J. Anim. Sci. 90:45–53. doi:10.2527/jas.2011-4031.

Thiruvenkadan, A. K., N. Kandasamy, and S. Panneerselvam. 2009. Inheritance of racing performance of trotter horses: An overview. Livest. Sci. 124:163–181. doi:10.1016/j.livsci.2009.01.010.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980.

Zhang, Z., J. Liu, X. Ding, P. Bijma, D. J. de Koning, and Q. Zhang. 2010. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. PLoS ONE 5(9):E12648. doi:10.1371/journal.pone.0012648.