American Genetic Association

Journal of Heredity, 2016, 537–543
doi:10.1093/jhered/esw042
Original Article
Advance Access publication August 3, 2016

OXFORD

Original Article

Heritability of Recurrent Exertional Rhabdomyolysis in Standardbred and Thoroughbred Racehorses Derived From SNP Genotyping Data

Elaine M. Norton, James R. Mickelson, Matthew M. Binns, Sarah C. Blott, Paul Caputo, Cajsa M. Isgren, Annette M. McCoy, Alison Moore, Richard J. Piercy, June E. Swinburne, Mark Vaudin, and Molly E. McCue

From the Veterinary Population Medicine Department, University of Minnesota, 225 Veterinary Medical Center, 1365 Gortner Ave, St. Paul, MN 55108 (Norton and McCue); Veterinary Biomedical Sciences Department, University of Minnesota, St. Paul, MN (Mickelson); Equine Analysis Systems, LLC, Midway, KY (Binns); School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK (Blott); 333 SW 14th Ave, Pompano Beach, FL (Caputo); School of Veterinary Science, University of Liverpool; Philip Leverhulme Equine Hospital, University of Liverpool Leahurst Campus, Wirral, UK (Isgren); Department of Veterinary Clinical Medicine, University of Illinois, Urbana, IL (McCoy); Moore Equine Services, Cambridge, Canada (Moore); Department of Clinical Sciences and Services, The Royal Veterinary College, Hertfordshire, UK (Piercy); Animal DNA Diagnostics LTD, Cambridge, UK (Swinburne); and Animal Health Trust, Suffolk, UK (Vaudin).

Address correpondence to Molly E. McCue at the address above, or e-mail: mccu0173@umn.edu.

Received April 1, 2016; First decision May 6, 2016; Accepted July 7, 2016.

Corresponding editor: Ernest Bailey

Abstract

Recurrent exertional rhabdomyolysis (RER) in Thoroughbred and Standardbred racehorses is characterized by episodes of muscle rigidity and cell damage that often recur upon strenuous exercise. The objective was to evaluate the importance of genetic factors in RER by obtaining an unbiased estimate of heritability in cohorts of unrelated Thoroughbred and Standardbred racehorses. Four hundred ninety-one Thoroughbred and 196 Standardbred racehorses were genotyped with the 54K or 74K SNP genotyping arrays. Heritability was calculated from genome-wide SNP data with a mixed linear and Bayesian model, utilizing the standard genetic relationship matrix (GRM). Both the mixed linear and Bayesian models estimated heritability of RER in Thoroughbreds to be approximately 0.34 and in Standardbred racehorses to be approximately 0.45 after adjusting for disease prevalence and sex. To account for potential differences in the genetic architecture of the underlying causal variants, heritability estimates were adjusted based on linkage disequilibrium weighted kinship matrix, minor allele frequency and variant effect size, yielding heritability estimates that ranged between 0.41–0.46 (Thoroughbreds) and 0.39–0.49 (Standardbreds). In conclusion, between 34–46% and 39–49% of the variance in RER susceptibility inThoroughbred and Standardbred racehorses, respectively, can be explained by the SNPs present

on these 2 genotyping arrays, indicating that RER is moderately heritable. These data provide further rationale for the investigation of genetic mutations associated with RER susceptibility.

Subject areas: Bioinformatics and computational genetics
Key words: muscle disease, RER, tying-up

Exertional rhabdomyolysis (ER) is a clinical syndrome in horses characterized by muscle cell damage, pain, stiffness, and cramping with exercise (Rossdale et al. 1985; MacLeay et al. 1999b). Recurrent exertional rhabdomyolysis (RER) is a specific form of ER, described in Thoroughbred and Standardbred racehorses, in which individuals have more than 1 clinical episode of ER while competing at an appropriate fitness level (Lentz et al. 1999; Isgren et al. 2010). Diagnosis of RER is primarily based on clinical signs and supportive evidence such as detection of elevated serum creatine kinase (CK) and aspartate aminotransferase (AST) activities after exercise. Muscle biopsy is predominantly utilized to rule-out other potential causes of ER, but may show nonspecific myopathic features such as centrally located nuclei and variable degrees of necrosis and regeneration (Valberg et al. 1993; Valberg et al. 1999). Estimated prevalence of RER is 6% in Standardbred racehorses and ranges from 5% to 10% in Thoroughbreds (MacLeay et al. 1999b; McGowan et al. 2002; Isgren et al. 2010). This relatively high prevalence has a significant financial impact on the racing industry, with an average of 5.8 consecutive training days lost per episode of RER in Thoroughbred racehorses (Jeffcott et al. 1982).

A number of studies have identified risk factors contributing to RER. In Thoroughbred racehorses, epidemiological studies have concluded that diet, exercise, sex, age, and stress are risk factors for development of clinical disease (MacLeay et al. 1999a; McGowan et al. 2002; Upjohn et al. 2005). In 2010, Isgren et al. identified similar risk factors in cases of RER in Standardbred racehorses, consistent with the hypothesis that RER in both breeds is phenotypically similar but also share the same risk factors. Pedigree analysis has provided strong evidence of an underlying genetic component in Thoroughbreds (MacLeay et al. 1999b; Dranchak et al. 2005; Oki et al. 2005), thus RER is a complex disease with both genetic and environmental risk factors contributing to phenotype. Recent genome wide association studies have identified statistically significant associations for RER in Thoroughbred racehorses (Tozaki et al. 2010; Fritz et al. 2012). However, the loci identified differ between studies and specific functional alleles and genes contributing susceptibility to RER are unknown.

Previous work estimated RER heritability based on pedigree data (Oki et al. 2005), but such estimates from pedigrees may be overestimated as a consequence of 1) small populations of highly related individuals often confounded by a shared environment; 2) assortative mating; 3) pedigree errors; and 4) ascertainment bias (Lee and Pollak 1997). Genotype data allows for a more precise estimate of narrow sense heritability by using data from a large group of unrelated individuals. Mixed linear models can be used to estimate heritability by fitting all SNPs simultaneously as random effects to provide an unbiased estimate of the variance explained by all SNPs (Yang et al. 2010). In contrast, the Bayesian model views all parameters as uncertain, and expresses this uncertainty over narrow sense heritability with a posterior distribution (Furlotte et al. 2014). The objective of this study was to estimate RER heritability in both Thoroughbred and Standardbred racehorses using whole genome SNP genotype data with different computational methods and parameters: including mixed linear and Bayesian models, utilizing standard and linkage disequilibrium weighted kinship matrixes.

Materials and Methods

Samples

Signalment, history and management data and whole blood samples for DNA isolation were collected from 245 case (174 females and 71 males) and 246 control (138 females and 108 males) Thoroughbred racehorses, and 107 case (63 females and 44 males) and 89 control (26 females and 63 males) Standardbred racehorses. A total of 72 Standardbreds and 405 Thoroughbred racehorses were from Europe, and 124 Standardbred and 86 Thoroughbred racehorses were from North America. Within the Thoroughbred cohort, ages ranged from 2 to 9 years old (mean 3 years) and in the Standardbred cohort ages ranged from 2 to 16 years (mean 5 years). Cases were defined as horses that had more than 1 episode of ER as reported by either the trainer or referring veterinarian. In 55.97% (197/352) of cases, measurement of elevated serum activities of CK and AST were available post-ER episode to further support the RER diagnosis. Control horses were defined as horses that did not have a history of ER observed by the trainer or referral veterinarian and had raced for at least 1 full season. Although RER can manifest in horses as young as 2 years of age; whenever possible, controls were selected from older horses that had been in training with the reporting trainer for more than 1 consecutive racing season to avoid a false negative diagnosis of RER.

Genotype Data

DNA was isolated from whole blood per manufacturer recommendations (Puregene Blood Core Kit, Qiagen). Thoroughbred horses were genotyped with the Illumina iSelect Equine SNP50 bead chip that contains 54 602 markers distributed across all 31 autosomes and the X-chromosome. Standardbred horses were genotyped with either the iSelect Equine SNP50 or the more recent 65 153 SNP chip (Illumina iSelect Equine SNP70 bead chip). Genotype imputation with haplotype phasing using BEAGLE software (Browning and Browning 2007; McCoy and McCue 2014) was performed to generate a uniform set of 73 689 SNP genotypes across the entire Standardbred cohort.

Quality control (QC) measures were performed on the genotyping data using the PLINK software package (Purcell et al. 2007). QC measures included tests for Hardy-Weinberg equilibrium, SNP and individual missingness and genotyping rates, discordant sex information and abnormally high heterozygosity (≥3 standard deviations [SDs] from the mean). All individuals passed QC. Individual SNPs that had a genotype success rate <90%, minor allele frequency (MAF) <1.0%, or deviated from Hardy-Weinberg equilibrium were eliminated from the dataset. After data pruning and exclusion of the sex chromosomes, a total of 45 447 SNPs and 61 101 autosomal SNPs remained in the dataset for analysis in the Thoroughbreds and Standardbreds, respectively.

The RER phenotype was treated as a binary trait (case/control). When pair-wise identity by descent estimations exceeded 0.25 (half sibling, grandparent-offspring); 1 individual per pair was randomly excluded from the analysis, yielding a total of 404 Thoroughbreds and 175 Standardbred racehorses for analysis. Sex for both breeds was included as a discrete covariate. Based on prior publications, prevalence of disease was estimated at 7.5% (published estimates ranged from 5% to 10%) for Thoroughbreds and 6% for Standardbreds (MacLeay et al. 1999b; McGowan et al. 2002; Isgren et al. 2010).

Analysis With a Mixed Linear Model

The Genome-wide Complex Trait Analysis (GCTA) software package (Yang et al. 2011) was used to estimate heritability from the SNP genotype data with a mixed linear model. A comprehensive analysis of the GCTA software tools, algorithms, and publications can be found online (http://cnsgenomics.com/software/gcta/, last accessed Janurary 2016). For a given phenotype, the output of GCTA provides an estimate of the genetic variance explained by all SNPs, the standard error (SE) of the estimate, and a likelihood ratio test of the alternative hypothesis (genetic variance ≠ 0) to that of the null hypothesis (genetic variance = 0) and the corresponding *P*-value (Yang et al. 2013).

Analysis With a Bayesian Model

Heritability was estimated from the SNP genotype data through a Bayesian posterior probability and the standard genetic relationship matrix (GRM) using code obtained from Furlotte, et al. (2014). A full description of the Bayesian model has previously been published (Furlotte et al. 2014). For a given phenotype, the output of the Bayesian model is the maximum posterior value for heritability and SD of the posterior probability distribution. The heritability estimates were then converted to a continuous liability scale using an ascertainment-corrected transformation as previously described (Lee et al. 2011).

Analysis With a Linkage Disequilibrium Adjusted Kinship Matrix

The software program Linkage Disequilibrium Adjusted Kinship (LDAK) (Speed et al. 2012) was used to create a weighted GRM (wGRM). LDAK assesses patterns of linkage disequilibrium (LD) through local pairwise correlation between SNPs. The extent to which a SNP signal is replicated by neighboring SNPs is calculated by summing the values in each row of the correlation matrix. Calculated weights are then determined so that the value of the SNP signal multiplied by the weight is equal to 1. The weighted values, or relative contribution of each SNP, are then utilized to create the wGRM. A comprehensive analysis of the LDAK algorithm can be found in Speed et al. (2012). Heritability estimates were then calculated with the wGRM using both GCTA and the Bayesian model.

Analysis of Sample Size on Heritability Estimates and SE

In order to assess the effect of sample size on heritability estimates and SE, heritability estimates were calculated using data from 200 horses from the Thoroughbred cohort to resemble the size of the Standardbred cohort, and the Standardbred cohort was reduced to 85 horses. For each breed, horses were randomly chosen, regardless of phenotype, using the software package R's random number generator without replacement (R Core Team 2014). Heritability estimates were calculated in GCTA using the standard GRM. This process was repeated a total of 100 times, after which the average of the heritability estimates and SE were calculated.

Analysis by Partitioning of SNPs

SNPs were partitioned into 5 bins with MAF boundaries of 0.1, 0.2, 0.3, 0.4, and 0.5, respectively as previously described (Lee et al. 2012). Standard GRMs and wGRM were constructed for each bin and heritability estimates were obtained using the –mgrm option in GCTA (Lee et al. 2013). SNPs were also partitioned into 2 groups with SNPs from autosomes 1 through 15 in the first group and SNPs from autosomes 16 to 31 in the second group. Heritability estimates were calculated for each group separately. The total heritability estimate was obtained by summing the estimates from both groups, which was then compared to the heritability estimate obtained from inclusion of all SNPs.

Results

Analysis With a Mixed Linear Model

Analysis of the SNP genotype data using GCTA with the software's default parameters and the standard GRM resulted in estimates of the genetic variance on the observed scale of 0.36 (SE = 0.13) for the Thoroughbred and 0.49 (SE = 0.30) for the Standardbred racehorses. Conversion of the estimates to a liability scale with the addition of disease prevalence resulted in heritability estimates of 0.39 (SE = 0.11) in the Thoroughbreds and 0.44 (SE = 0.27) in the Standardbreds. When sex was added to the analysis as a covariate, the heritability estimate decreased to 0.34 (SE = 0.12) in the Thoroughbreds and to 0.45 (SE = 0.27) in the Standardbreds (Table 1). All estimates resulted in statistically significant *P*-values.

Analysis With a Bayesian Model

Heritability estimates using the Bayesian model and the standard GRM resulted in a posterior probability of 0.36 (SD = 0.12) in the

Table 1. Heritability estimates using both mixed linear model (GCTA) and Bayesian model (Bayes) with the standard GRM and the weighted GRM (wGRM) when both disease prevalence and sex were taken into account, and for GCTA when using an adjusted scale parameter (S=0) to account for the uncertainty around minor allele frequency and variant effect size

	GCTA GRM	Bayes GRM	GCTA wGRM	Bayes wGRM	S=0 GRM	S=0 wGRM
TB cohort						
h^2	0.34	0.34	0.42	0.42	0.41	0.46
SE/SD	0.12	0.11	0.12	0.11	0.12	0.14
P-value	2.320e-04		1.010e-05		3.111e-05	4.320e-05
STDB cohort						
h^2	0.45	0.45	0.39	0.39	0.45	0.41
SE/SD	0.27	0.20	0.26	0.20	0.27	0.27
P-value	3.563e-03		1.786e-02		3.530e-03	1.801e-02

Thoroughbreds and 0.49 (SD = 0.22) in the Standardbreds. Addition of sex as a covariate did not greatly alter the heritability estimates, with 0.36 (SD = 0.12) and 0.50 (SD = 0.22) in the Thoroughbreds and Standardbreds, respectively. Transformation of the data from the observed scale to a liability scale via the addition of disease prevalence resulted in a heritability estimate of 0.34 (SD = 0.11) in the Thoroughbreds and 0.45 (SD = 0.20) in the Standardbreds (Table 1).

Analysis With an LDAK

Analysis using GCTA with the substitution of the wGRM obtained from LDAK, adjusting for both disease prevalence and sex, resulted in a heritability estimate of 0.42 (SE = 0.12) in the Thoroughbred and 0.39 (SE = 0.26) in the Standardbred cohort (Table 1). Analysis with the Bayesian model using the wGRM resulted in a heritability estimate of 0.42 (SD = 0.11) in the Thoroughbreds and 0.39 (SD = 0.20) in the Standardbreds (Table 1).

Analysis of Sample Size on Heritability Estimates and SE

In order to assess the effect of sample size on the heritability estimates and SEs, GCTA analysis with the inclusion of sex and disease prevalence was repeated 100 times in randomly selected groups of 200 horses from the Thoroughbred cohort and 85 horses from the Standardbred cohort. The mean heritability estimate from the resultant 100 analyses was 0.43 with a mean SE of 0.24 in the Thoroughbreds as compared to the original estimate of 0.34 with a SE: 0.12. The mean heritability estimate for the Standardbred data was 0.43 with a mean SE of 0.39, as compared to the original estimate of 0.45 with a SE of 0.27; however, 60% of the estimates in the Standardbreds resulted in nonsignificant *P*-values.

Analysis of MAF and Variant Effect Size

Heritability estimates from dense genotyping data may be overestimated when the assumed relationship between the causal variant's MAF and effect size, represented by a scale parameter, does not reflect the data's genotypic architecture (Speed et al. 2012; Lee et al. 2013). Heritability estimates were repeated with an adjusted scale parameter by using the –grm-adj function in GCTA, resulting in an estimate of 0.41 (SE = 0.12) in the Thoroughbreds and 0.45 (SE = 0.27) in the Standardbreds (Table 1), as compared to 0.34 and 0.45 observed in the original analysis. Using the wGRM and adjusted scale parameter resulted in heritability estimates of 0.46 (SE = 0.14) for the Thoroughbreds and 0.41 (SE = 0.27) for the Standardbreds (Table 1), as compared to 0.42 and 0.39 in the original analysis.

A MAF-stratification approach has also been recommended to account for the uncertainty around the underlying genetic architecture (Lee et al. 2012; Lee et al. 2013). Using this approach, the sum of the heritability estimates from the standard GRMs were 0.46 in the Thoroughbreds and 0.49 in the Standardbreds. Notably, SE was markedly higher per bin in the Standardbred cohort (Table 2).

Analysis of Population Stratification and Cryptic Relatedness

To evaluate the presence of population stratification, the total heritability estimate obtained from summing the results from partitioning across chromosomal groups was compared to the estimate obtained from inclusion of all SNPs (Speed et al. 2012). For the Thoroughbred cohort, the total variance explained was 0.49 when the chromosomes were fitted separately in groups compared to a total of 0.34 when the chromosomes were fitted together, yielding a difference of 0.15. For the Standardbred cohort, the total variance explained was 0.60 when the chromosomes were fitted separately by groups compared to a total of 0.45 when the chromosomes were fitted together, yielding a difference of 0.15.

To evaluate the effect of cryptic relatedness on the heritability estimates, 10% of the population was randomly removed from the dataset 100 times and heritability estimates were recalculated using GCTA with both the standard GRM and the wGRM. For the Thoroughbred cohort, the mean heritability estimate was 0.36 for the standard GRM and 0.46 for the wGRM. For the Standardbred cohort, the mean heritability estimate was 0.43 for the standard GRM and 0.42 for the wGRM. All estimates were consistent with the estimates obtained from their respective analyses using the full dataset.

Discussion

Traditionally, heritability estimates have been obtained using information from pedigree data. However, the recent availability of SNP arrays for many species now allows heritability to be estimated from SNP genotype data obtained from unrelated or distantly related individuals. In this report we calculated heritability of RER in populations of 404 Thoroughbred and 175 Standardbred horses using more than 45 000 (Thoroughbreds) and 61 000 (Standardbreds) SNPs with 2 different computational methods (mixed linear and Bayesian models) whilst accounting for genetic relatedness. Both computational models estimate heritability based on a comparison of overall similarities in phenotype between pairs of individuals with their total SNP similarities, providing an overall estimate of the genetic influence on a trait (Viding et al. 2013). After accounting for disease prevalence and potential confounding factors such as sex and cryptic relatedness, heritability estimates for RER ranged between 0.34–0.46 in Thoroughbred and 0.39–0.49 in Standardbred racehorses. Heritability estimates obtained using either the mixed linear model or the Bayesian model were not remarkably different.

Table 2. Heritability estimates using GCTA and a MAF stratification approach with the standard GRM when both disease prevalence and sex were taken into account

	Bin1: MAF < 0.1	Bin 2: MAF 0.1–0.2	Bin 3: MAF 0.2–0.3	Bin 4: MAF 0.3–0.4	Bin 5: MAF 0.4–0.5	Sum
TB cohort GRM						
Heritability	0.02	0.04	0.00	0.40	0.00	0.46
SE	0.01	0.16	0.18	0.20	0.16	
STDB cohort GRM						
Heritability	0.27	0.00	0.00	0.00	0.22	0.49
SE	0.25	0.38	0.39	0.40	0.39	

SNPs were divided into bins based on MAF.

Heritability estimates are highly dependent on the represented population since the effects of environmental variance and additive and non-additive genetic variances are population-specific. Furthermore, the accuracy of heritability estimates is dependent on the pedigree structure and introduction of unaccounted-for bias. Limited information or pedigree errors within a highly inbred population, assortative mating, and ascertainment bias (selection of pedigrees that have a high proportion of affected individuals) can all lead to overestimation of heritability with larger SEs (Lee and Pollak 1997). Furthermore, to achieve an unbiased estimate of genetic variance, the data must be representative of the general population and include all potential confounders (Visscher et al. 2008).

Our estimates of heritability address these biases by 1) using SNP-genotype based approaches which do not depend on pedigree data; 2) including a large population of individuals with wide range of genetic backgrounds and across multiple farms; and 3) accounting for known confounders. Sex was included as a confounder based on results of studies in both Thoroughbred and Standardbred racehorses indicating that females are at a higher risk of developing RER (MacLeay et al. 1999a; McGowan et al. 2002; Upjohn et al. 2005; Isgren et al. 2010). The addition of sex decreased the heritability estimates in both breeds, indicating that at least a portion of the phenotypic variance between the cases and controls could be explained by sex. Age has also been determined to be a risk factor for RER in Thoroughbred racehorses but was not determined to be a risk factor in Standardbred racehorses by Isgren et al. (2010). The difference in age as a risk factor between these breeds may be a reflection of the average career length of Thoroughbred versus Standardbred racehorses (Isgren et al. 2010). In our dataset, age at the time of sampling was available for all horses; however, the age of RER onset was not available for most cases. Age was included as a quantitative covariate in both the mixed linear and Bayesian model and had minimal effect on the overall heritability estimates (data not shown). However, we did not include age in the final reported heritability estimate models since age at sampling was not considered a true factor affecting disease risk.

The large SEs for the Standardbred heritability estimates likely reflect the relatively small number of horses in the Standardbred cohort compared to the Thoroughbred cohort. To test this theory, 200 horses were randomly chosen 100 times from the Thoroughbred dataset for GCTA analysis using the standard GRM. The mean heritability estimate for the data was 0.43 with a mean SE of 0.24, as compared to the original estimate of 0.34 with a SE of 0.12. Therefore, it is likely that the addition of more Standardbred horses would decrease the SE and result in a more precise estimate of heritability in this breed. This is similar to findings from Yang et al. who estimated the heritability of height in humans using GCTA. In that study, the investigators determined that the average heritability estimates across all SNPs were not dependent on sample size but that sampling error increased as the population size decreased (Yang et al. 2010). Similarly, 85 horses were randomly chosen 100 times from the Standardbred cohort and heritability was estimated using GCTA with the standard GRM. The mean heritability estimate for the data was 0.43 with a mean SE of 0.39, as compared to the original estimate of 0.45 with a SE of 0.27. However, 60% of these estimates resulted in nonsignificant *P*-values, indicating that reducing the Standardbred cohort by half would have insufficient power to estimate heritability.

A pedigree-based estimate of RER heritability in a cohort of Japanese Thoroughbred racehorses trained and stabled at 1 of 2 facilities was approximately 0.42 (Oki et al. 2005). This heritability was higher than the 0.34 estimated in our Thoroughbred cohort when taking into account sex and disease prevalence but not taking into account potential differences in genetic architecture. It is perhaps not surprising that the estimates in the current study using Thoroughbred racehorses across multiple environments were for the most part lower than those in the Japanese study, as heritability estimates from a single population residing on the same farm might be overestimated due to shared genetics and shared environment. However, the most precise estimates of heritability should be obtained when accounting for the genetic architecture of the underlying causal variants (Speed et al. 2012; Lee et al. 2013; Speed et al. 2013).

Without prior knowledge of the causal variants for RER, it is impossible to determine which genetic architecture would be the most appropriate model. Therefore, 3 approaches were used to adjust heritability estimates based on LD, MAF, and variant effect size. It has been proposed that heritability estimates using genotype data are sensitive to blocks of SNPs in LD, where SNPs in high LD with a causal variant are overrepresented leading to inflated estimates of heritability (Speed et al. 2012). The tagging of causal SNPs in blocks of uneven LD might account for the difference in heritability estimates obtained using the standard GRM versus the wGRM. By adjusting for LD through a wGRM, causal SNPs are theoretically more evenly weighed. In our dataset, the heritability estimate decreased in the Standardbred cohort to 0.39 and increased in the Thoroughbred cohort to 0.42 (both GCTA and Bayesian models). In the Standardbred cohort, causal SNPs within blocks of high LD might have been overrepresented, leading to inflation of the heritability estimate; whereas in the Thoroughbred cohort casual SNPs within areas of low LD may have been underrepresented decreasing the heritability estimate (Speed et al. 2012; Lee et al. 2013).

By default, GCTA scales the heritability estimate based on heterozygosity across the genome (Lee et al 2013). However, Speed et al. (2012) proposed varying the scale parameter to assume independence between effect size and MAF. In our study population, alteration of the scale parameter resulted in similar estimates of heritability in the Standardbred cohort but higher estimates of heritability in the Thoroughbred cohort. Alternatively, Lee et al. (2013) recommended accounting for the relationship between MAF and variant effect size with a MAF-stratification approach. In this approach, SNPs are partitioned into bins based on MAF and heritability estimates are calculated for each bin. For the Standardbred cohort, the largest estimates were obtained for the bin with SNPs at the lowest MAF (bin 1: MAF < 0.1) and highest MAF (bin 5: MAF 0.4–0.5); whereas, for the Thoroughbred cohort the bin with the largest estimate was bin 4 (MAF 0.3–0.4). This may explain the difference in heritability estimates between varying the scale parameter and using the MAF binning approach.

In both cohorts, manipulation of the data either increased or decreased the heritability estimates from what was obtained when using the default parameters for the mixed linear or Bayesian model. However, without knowing the causal variants for RER, it is impossible to determine which model is the most appropriate. Nevertheless, all estimates were consistent with a trait of moderate heritability.

The presence of population stratification, genetically distinct groups, or cryptic relatedness, within a dataset may confound heritability estimates. To assess the presence of population stratification, SNPs can be partitioned into groups based on chromosome and the total summed across groups. Summed values will be greater than heritability estimates obtained from inclusion of all SNPs when SNPs on different chromosomes are correlated more than expected

by chance, indicating population substructure (Yang et al. 2013). In our analysis, the difference in total variance for all 31 autosomes analyzed individually versus jointly was 0.15 higher in both the Thoroughbred and Standardbred cohorts, indicating some influence from population stratification. To assess the effect of cryptic relatedness, or a few related individuals influencing the heritability estimate, 10% of the population was randomly removed from the analysis for each cohort and heritability estimates repeated. The mean heritability across all 100 replicates was similar to that obtained from the whole population for both the Thoroughbreds and Standardbreds, indicating little evidence that cryptic relatedness is biasing the heritability estimates in this analysis.

SNP-based heritability estimates are limited to the genetic variability that can be explained by the common SNPs present on the genotyping arrays, but cannot account for causal variants that are not inherited together (in LD) with these SNPs; nor can it include other genetic variations contributing to the disease phenotype such as insertions, deletions or copy number variants (Manolio et al. 2009; Lee et al. 2011; Yang et al. 2011; Lee et al. 2012). Completely unbiased estimates of heritability can only be obtained if all of the causal genetic variants are represented. Therefore, SNP based heritability estimates may provide a lower-limit estimate for narrow sense heritability (Manolio et al. 2009; Lee et al. 2011; Yang et al. 2011; Lee et al. 2012). The unequal number of SNPs used for the Thoroughbred and Standardbred analyses (45 447 and 61 101 SNPs, respectively) prevents direct comparison of the estimates between breeds as the additional 15 000 SNPs in the Standardbred cohort may capture a larger proportion of the genetic variance leading to a higher heritability estimate. However, based on long blocks of LD in both breeds, we do not suspect that the heritability estimates would vary significantly with this relatively small change in SNP density (McCue et al. 2012).

Nonetheless, the results indicate that 34–46% of the heritability of RER in Thoroughbred racehorses and 39–49% of the heritability in Standardbred racehorses can be explained by their respective SNP data, supporting the conclusion that RER is moderately heritable. Previous studies have attempted to identify RER susceptibility loci through whole genome scanning methods with microsatellites (Tozaki et al. 2010) or SNPs (Fritz et al. 2012). However, those studies identified different potential risk loci, with moderate or ambiguous statistical support, and our current heritability data suggests that larger sample sizes and higher SNP marker density than were available in these previous studies will be necessary to identify RER susceptibility loci.

In conclusion, SNP based heritability estimates for RER in Thoroughbred and Standardbred racehorses indicate that 34–46% and 39–49%, respectively, of genetic variance for RER can be explained by the common SNPs present on the current genotyping arrays. It is important to note that heritability estimates do not provide information on the number of genes involved, the interaction or penetrance of these genes, nor the mode by which these genes are inherited. Heritability estimates are also sensitive to small datasets, population substructure, and the data's genotypic architecture. However, heritability estimates do provide valuable insight into the genetic contribution of a complex disease. The results presented here supports the conclusion that RER is moderately heritable in both Standardbred and Thoroughbred racehorses. This has important implications for the racing industries, as it is likely that RER affected horses of both breeds and sexes are capable of passing susceptibility genes on to their offspring.

Funding

National Institutes of Health (1K081R055713-01A2 to M.E.M.; T32 Institutional Training Grant in Comparative Medicine and Pathology, 5T320D010993-12, supporting E.M.N.); the University of Minnesota Wellstone Muscular Dystrophy Center (P30 Pilot funding); and the Horserace Betting Levy Board.

Acknowledgments

The authors would like to acknowledge Dr Stephanie Valberg for her contribution of the 86 North American Thoroughbred samples from her Equine Neuromuscular Disease Laboratory repository, and the numerous owners, trainers, and referring veterinarians who participated in the sampling of horses for this study.

References

Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 81:1084–1097.

Dranchak PK, Valberg SJ, Onan GW, Gallant EM, MacLeay JM, McKenzie EC, De La Corte FD, Ekenstedt K, Mickelson JR. 2005. Inheritance of recurrent exertional rhabdomyolysis in thoroughbreds. *J Am Vet Med Assoc*. 227:762–767.

Fritz KL, McCue ME, Valberg SJ, Rendahl AK, Mickelson JR. 2012. Genetic mapping of recurrent exertional rhabdomyolysis in a population of North American Thoroughbreds. *Anim Genet*. 43:730–738.

Furlotte NA, Heckerman D, Lippert C. 2014. Quantifying the uncertainty in heritability. *J Hum Genet*. 59:269–275.

Isgren CM, Upjohn MM, Fernandez-Fuente M, Massey C, Pollott G, Verheyen KL, Piercy RJ. 2010. Epidemiology of exertional rhabdomyolysis susceptibility in standardbred horses reveals associated risk factors and underlying enhanced performance. *PLoS One*. 5:e11594.

Jeffcott LB, Rossdale PD, Freestone J, Frank CJ, Towers-Clark PF. 1982. An assessment of wastage in thoroughbred racing from conception to 4 years of age. *Equine Vet J*. 14:185–198.

Lee C, Pollak EJ. 1997. Influence of sire misidentification on sire x year interaction variance and direct-maternal genetic covariance for weaning weight in beef cattle. *J Anim Sci*. 75:2858–2863.

Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME, Keller MC, Visscher PM, Wray NR. 2012. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet*. 44:247–250.

Lee SH, Wray NR, Goddard ME, Visscher PM. 2011. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet*. 88:294–305.

Lee SH, Yang J, Chen GB, Ripke S, Stahl EA, Hultman CM, Sklar P, Visscher PM, Sullivan PF, Goddard ME, *et al*. 2013. Estimation of SNP heritability from dense genotype data. *Am J Hum Genet*. 93:1151–1155.

Lentz LR, Valberg SJ, Balog EM, Mickelson JR, Gallant EM. 1999. Abnormal regulation of muscle contraction in horses with recurrent exertional rhabdomyolysis. *Am J Vet Res*. 60:992–999.

MacLeay JM, Sorum SA, Valberg SJ, Marsh WE, Sorum MD. 1999a. Epidemiologic analysis of factors influencing exertional rhabdomyolysis in Thoroughbreds. *Am J Vet Res*. 60:1562–1566.

MacLeay JM, Valberg SJ, Sorum SA, Sorum MD, Kassube T, Santschi EM, Mickelson JR, Geyer CJ. 1999b. Heritability of recurrent exertional rhabdomyolysis in Thoroughbred racehorses. *Am J Vet Res*. 60:250–256.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, *et al*. 2009. Finding the missing heritability of complex diseases. *Nature*. 461:747–753.

McCoy AM, McCue ME. 2014. Validation of imputation between equine genotyping arrays. *Anim Genet*. 45:153.

McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guérin G, Hasegawa T, Hill EW, *et al*. 2012. A high density SNP array for

the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet*. 8:e1002451.

McGowan CM, Fordham T, Christley RM. 2002. Incidence and risk factors for exertional rhabdomyolysis in thoroughbred racehorses in the United Kingdom. *Vet Rec*. 151:623–626.

Oki H, Miyake T, Hasegawa T, Sasaki Y. 2005. Estimation of heritability for Tying-up syndrome in the Thoroughbred racehorse by Gibbs sampling. *J Anim Breed Genet*. 122:289–293.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al*. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 81:559–575.

R Core Team. 2014. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna (Austria): R Foundation for Statistical Computing. Available from: URL http://www.R-project.org/, last accessed January 2016.

Rossdale PD, Hopes R, Digby NJ, offord K. 1985. Epidemiological study of wastage among racehorses 1982 and 1983. *Vet Rec*. 116:66–69.

Speed D, Hemani G, Johnson MR, Balding DJ. 2012. Improved heritability estimation from genome-wide SNPs. *Am J Hum Genet*. 91:1011–1021.

Speed D, Hemani G, Johnson MR, Balding DJ. 2013. Response to Lee *et al*.: SNP-based heritability analysis with dense data. *Am J Hum Genet*. 93:1155–1157.

Tozaki T, Hirota K, Sugita S, Ishida N, Miyake T, Oki H, Hasegawa T. 2010. A genome-wide scan for tying-up syndrome in Japanese Thoroughbreds. *Anim Genet*. 41(suppl 2):80–86.

Upjohn MM, Archer RM, Christley RM, McGowan CM. 2005. Incidence and risk factors associated with exertional rhabdomyolysis syndrome in National Hunt racehorses in Great Britain. *Vet Rec*. 156:763–766.

Valberg S, Jönsson L, Lindholm A, Holmgren N. 1993. Muscle histopathology and plasma aspartate aminotransferase, creatine kinase and myoglobin changes with exercise in horses with recurrent exertional rhabdomyolysis. *Equine Vet J*. 25:11–16.

Valberg SJ, Mickelson JR, Gallant EM, MacLeay JM, Lentz L, de la Corte F. 1999. Exertional rhabdomyolysis in quarter horses and thoroughbreds: one syndrome, multiple aetiologies. *Equine Vet J*. 31(suppl 30):533–538.

Viding E, Price TS, Jaffee SR, Trzaskowski M, Davis OS, Meaburn EL, Haworth CM, Plomin R. 2013. Genetics of callous-unemotional behavior in children. *PLoS One*. 8:e65789.

Visscher PM, Hill WG, Wray NR. 2008. Heritability in the genomics era–concepts and misconceptions. *Nat Rev Genet*. 9:255–266.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, *et al*. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 42:565–569.

Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 88:76–82.

Yang J, Lee SH, Goddard ME, Visscher PM. 2013. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol Biol*. 1019:215–236.