ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi:10.1111/j.1365-2052.2010.02127.x

Morphological variation in the horse: defining complex traits of body size and shape

S. A. Brooks*, S. Makvandi-Nejad[†], E. Chu[†], J. J. Allen[†], C. Streeter*, E. Gu[†], B. McCleery[†], B. A. Murphy[‡], R. Bellone[§] and N. B. Sutter[†]

*Department of Animal Science, College of Agriculture and Life Sciences, Cornell University, Ithaca, NY 14853, USA. [†]Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA. [‡]School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland. [§]Department of Biology, University of Tampa, Tampa, FL 33606, USA

Summary

Horses, like many domesticated species, have been selected for broad variation in skeletal size. This variation is not only an interesting model of rapid evolutionary change during domestication, but is also directly applicable to the horse industry. Breeders select for complex traits like body size and skeletal conformation to improve marketability, function, soundness and performance in the show ring. Using a well-defined set of 35 measurements, we have identified and quantified skeletal variation in the horse species. We collected measurements from 1215 horses representing 65 breeds of diverse conformation such as the American Miniature, Shetland Pony, Arabian Horse, Thoroughbred, Shire and Clydesdale. Principal components analysis has identified two key dimensions of skeletal variation in the horse. Principal component 1 is positively correlated with every measurement and quantifies overall body size. Principal component 2 captures a pattern of bone widths vs. lengths and thus quantifies variation in overall bone thickness. By defining these complex skeletal traits, we have created a framework for whole genome association studies to identify quantitative trait loci that contribute to this variation.

Keywords complex trait, horse, morphology, principal components analysis, skeleton.

Introduction

Selective modification of domesticated plants and animals began millennia ago and is still ongoing today. In the present era of continually decreasing costs for whole genome genotyping and sequencing, it is now feasible to work in any number of species. Domesticated mammals are particularly powerful systems for gene discovery, as well as valuable models for identifying traits under selection. To fully leverage the intrinsic genetic value of these highly structured populations, however, it is critical to obtain high-quality phenotypes for study. The horse, like other domesticated species, has been moulded through selection within breeds into diverse skeletal morphologic types, often to fit specific functions. As a result of this selection, there are now breeds of horse that greatly differ from one another in body size and shape.

Address for correspondence

N. B. Sutter, C3-179 Veterinary Medical Center, Cornell University, Ithaca, NY 14853, USA.
Email: nbs39@cornell.edu

Accepted for publication 20 September 2010

Quantitative measurement of the horse's morphology has been previously carried out by a number of groups (Sadek *et al.* 2006; Weller *et al.* 2006; Komosa & Purzyc 2009). In each case, only one or two breeds were studied, and therefore the identified patterns of variation were primarily within breeds and not between them. High estimates of heritability for skeletal measures have been calculated in several breeds of horse. In the Murgese, for example, heritability of cannon bone circumference was estimated at 0.44 (Dario *et al.* 2006). However, most of the previous studies documenting the heritability of skeletal traits have focused on qualitatively described traits like 'over at the knee,' 'turned-in feet,' 'weak pasterns' and 'tied below the knee', instead of quantitative traits as described in this study (Love *et al.* 2006, Dario *et al.* 2006). Here, we aim to define skeletal variation for the horse species as a whole by collecting quantitative body measurements from a broad set of horse breeds with diverse body types. We have used a principal components analysis (PCA) to identify patterns of variation across 35 measurements of the head, neck, trunk and limbs. Unlike individual measurements, PCA can identify whole-body patterns like bone thickness. Furthermore, the distillation of complex skeletal morphometric data sets through PCA has proven successful in other domestic mammals such as the dog (Sutter *et al.* 2007). These data will enable mapping of genes contributing to these complex traits.

Materials and methods

Collection of measurements

We developed a protocol for body measurement of live horses using a provided tape measure and a uniform set of illustrated instructions (Fig. S1). A laboratory staff member, collaborator, or individual horse owner used the protocol to guide measurement of each horse and to subjectively score a set of traits coded as factors for body condition score (Henneke *et al.* 1983), muzzle profile ('dish-faced' to neutral to Roman nosed), degree of feathering (none to copious), bite evenness (underbite to neutral to overbite) and bone thickness (Fig. S2a–f). We noted if the horse was shod on the fore or hind hooves. Vital statistics and basic history (date or year of birth, registered name, registry, barn name, brief notes on any injury or disease, a photo and a pedigree for each horse) was also collected from the horse owner or agent. The coefficient of variation for each measure is shown in Fig. S3. All of the horses measured were at least four years of age at the time of trait collection and therefore relatively skeletally mature (range 4–34.5 years, median age 11.5; Fig. S2a) (Sadek *et al.* 2006).

For each horse, we collected the following quantitative measurements: (1) Eye to eye width: standing in front of the horse, measure across the forehead between the inside corners of the eyes. Keep the tape taut and straight. (2) Jaw width: across the underside of the head, measure straight across from the outsides of the cheekbones. Keep the tape taut and straight. (3) Head length: starting between the top corners of the two nostrils, measure straight to the front of the poll. (4) Muzzle circumference: take the circumference of the muzzle underneath the halter, settling the tape directly in front of the cheekbones. (5) Left eye to mouth length: measure from the corner of the mouth to the back corner of the eye. (6) Left eye to jaw length: measure from the back corner of the eye to the deepest point of the cheek curve. (7) Left ear length: on the side of the ear closest to the poll, measure from the base to the tip of the ear. (8) Neck length, head level with withers: measure from the poll to the withers, with your horse's head level with the withers. (9) Neck length, head down to the ground: measure as in #8, but with your horse's head stretched as close to the ground as possible. (10) Neck circumference at throat latch: settle the tape where the throat latch of the bridle goes. Pull it snugly but not uncomfortably for your horse. (11) Neck circumference at base: settle the tape just in front of the withers. Let it rest on the chest and curve comfortably around the base of the neck. (12) Height at withers: measure from the ground to the highest point of the withers. Keep the tape taut and straight. (13) Height at croup: measure from the ground straight up to the highest point of the rump. (14) Height at dock: measure from the ground straight up to the base of the tail. (15) Tail length: start at the end of the bony portion of the tail and measure to the dock. (16) Withers to croup, straight tape: measure across the back from the withers to the point of the croup. The tape will not touch the back except at the ends. (17) Withers to croup, contoured: measure the back from the withers to the croup, allowing the tape to relax and touch the entire length of the spine. (18) Length from croup to dock: measure from the point of the croup to the base of the tail. (19) Chest width: feel for the humeral bones that project forward out of the chest and measure the distance between the outside edges. Keep the tape taut and straight. (20) Barrel girth at heart: settle the measuring tape where the girth of the saddle fits, directly behind the forelegs. (21) Barrel girth, maximum: measure around the barrel at greatest circumference. (22) Left forearm length: measure from the point of the elbow to the back of the kneecap. (23) Left fore cannon length: measure from the back of the knee cap to the ergot. (24) Left fore cannon mid-point circumference: measure around the cannon bone halfway between the knee and the ergot. (25) Left fore pastern length: measure from the bottom of the fetlock joint to the top of the coronet. (26) Left fore pastern circumference: measure around the pastern. (27) Left fore coronet circumference: measure around the coronet. (28) Left hoof length: measure from the coronet to the bottom of the hoof. (29) Left gaskin length: measure from the stifle joint to the point of the hock. Keep the tape taut and straight. (30) Left hind cannon length: measure from the point of the hock to the ergot (bottom of the fetlock). For measurements 31–35 of the hind limb (left hind cannon mid-point circumference, left hind pastern length, left hind pastern circumference, left hind coronet circumference and left hind hoof length), the same directions were followed as for measurements 24–28.

Body measures were recorded in units of inches, as the majority of owners were from within the United States and most familiar with this system of units. Results reported here have been converted to SI units. Horses were measured while standing on level, solid ground and restrained by either cross-ties or a halter and rope. Fewer than ~2% of horses (<20) were found to be non-compliant with the measurement protocol upon attempted measurement. These horses were not included in the study. Our sample collection protocol is approved by the Cornell University Institutional Animal Care and Use Committee as protocol number 2007-0155.

Breeds

Sample collection was targeted to horse breeds with extreme large or small body sizes, thick or thin bones or diverse historical origins. The number of horses collected by breed and sex are listed in Table S1. The total collection for each breed is as follows: 6 Akhal-Teke, 16 American Belgian, 45 American Miniature, 9 Andalusian, 31 Appaloosa, 13 Appendix, 50 Arab, 4 Ardennais, 1 Bashkir, 1 Belgian Warmblood, 3 Brabant, 4 Caspian, 1 Chincoteague Pony, 2 Cleveland Bay, 25 Clydesdale, 2 Colorado Ranger Horse, 25 Connemara Pony, 3 Dartmoor Pony, 12 Draft Light Horse Cross, 18 Dutch Warmblood, 8 Falabella, 1 Fell Pony, 47 Fjord, 10 Friesian, 1 Gypsy Vanner, 8 Haflinger, 10 Half Arab, 10 Hanoverian, 5 Holsteiner, 28 Icelandic Horse, 1 Irish Draught Horse, 8 Irish Sport Horse, 3 Lipizzan, 3 Lusitano, 7 Missouri Fox Trotter, 44 Morgan, 8 Mountain Horse, 18 Mustang, 7 N. Am. Spanish Colonial, 7 National Show Horse, 8 Oldenburg, 45 Paint, 60 Paso Fino, 29 Percheron, 1 Peruvian Paso, 1 Pinto Horse, 15 Polo Pony, 2 Pony of the Americas, 114 Quarter Horse, 13 Saddlebred, 1 Selle Francais, 12 Shetland Pony, 28 Shire, 8 Sport Horse, 54 Standardbred, 6 Suffolk Punch, 12 Swedish Warmblood, 3 Tarpan, 30 Tennessee Walking Horse, 221 Thoroughbred, 23 Trakehner, 3 Welsh Cob (D), 4 Welsh Mountain Pony (A), 16 Welsh Pony (B) and 1 Westfalen. The total was 1215 horses.

Data handling and statistics

Sample and trait data were initially captured onto paper forms. Data were entered to a database and double-checked by a second lab member. After collecting 35 measurements from each horse, we excluded two measures from further analysis. These measures were tail length, which is inconsistent in some breeds because of docking, and barrel girth, which is confounded by a horse's body condition and in mares, pregnancy. Extreme outlier measurements were identified by calculating the median and interquartile range for each measurement in each breed with at least three samples. We then removed from the dataset 15 of the 22 horses that had at least one measurement value greater than ten times the interquartile range from the median value for that measurement in that breed (two Am. Miniature, one Appaloosa, three Connemara Pony, two Quarter Horse, one Standardbred, one Suffolk Punch, four Tenn. Walking Horse and one Thoroughbred). The remaining seven of the 22 horses (one each of Akhal-Teke, Fjord, Hanoverian, Irish Sport Horse, N.A. Spanish Colonial, Oldenburg and Paso Fino) with measures beyond this range were subjectively retained, because their breeds had extremely small interquartile ranges (often because of highly similar measurement values in relatively small sample sizes). We removed an additional six horses with measures that were between five and ten times the interquartile range from the median for their breed (one each of Lipizzan, Oldenburg, Standardbred, Trakehner and two Falabella). This was based on subjective assessment that the measurements were unrealistic outliers (gross errors in measurement or recording). We ignored a horse's breed or sex when making a determination that it would be excluded from the dataset because of outlier measurements. Statistical analysis was conducted using R v2.10.1, PERL v5.10.1 and JMP v8.0 software (SAS Institute Inc.).

We conducted PCA on the remaining 33 quantitative measures (excluding tail length and barrel girth) for the set of 1215 horses that met our inclusion criteria. We used the correlation matrix rather than covariance matrix because of the large variation in scale between our different measures (e.g. median heart girth is 1.9 m but median eye width is only 18 cm).

Results

We aim to identify the patterns of skeletal size and shape variation that exist in the domestic horse species. Our goal here is to derive robust and quantifiable morphologic traits that can subsequently be subjected to genetic analysis for discovery of genes contributing to these traits. We therefore collected 35 body measurements from the head, neck, trunk and limbs using the most readily identifiable body landmarks (Table S1).

To examine body size variation in the horse, we first looked at variation in the height at the withers across the horse species and then within specific breed populations. The extremes for height in our dataset are an American Miniature mare and stallion, with recorded withers heights of 74 and 76 cm, respectively, and two 2- m-tall Shire and Clydesdale stallions. We therefore observe over a 2.5-fold range in withers height in horses, irrespective of sex. The median withers height for the 1215 horses is 1.6 m, and half of all measured horses are between 1.5 and 1.6 m at the withers (Fig. 1). Despite the large variation in the species as a whole, each breed is characterized by a much narrower range of heights (Fig. 1). For example, our largest sample collection from a single breed is from Thoroughbreds (n = 221), where the shortest is 1.3 m and the tallest is 1.8 m. Thus, the height range within this breed is only 1.4-fold, far smaller than the range for the species as a whole.

To identify the broad patterns of variation in body size and shape in the horse species, we subjected our set of 33 quantitative measurements to a PCA. We examined the first and second PCs in detail. Although subsequent PCs may contain biologically relevant patterns, we focused on just the first two PCs for two reasons. First, inspection of the scree plot of the proportion of variance explained by each PC shows a levelling off for principal component 3 (PC3) and onward (Fig. S4). Second, we also applied the commonly used 'stopping rule' stating that PCs retained for analysis should explain a greater proportion of variance than $1/n$,

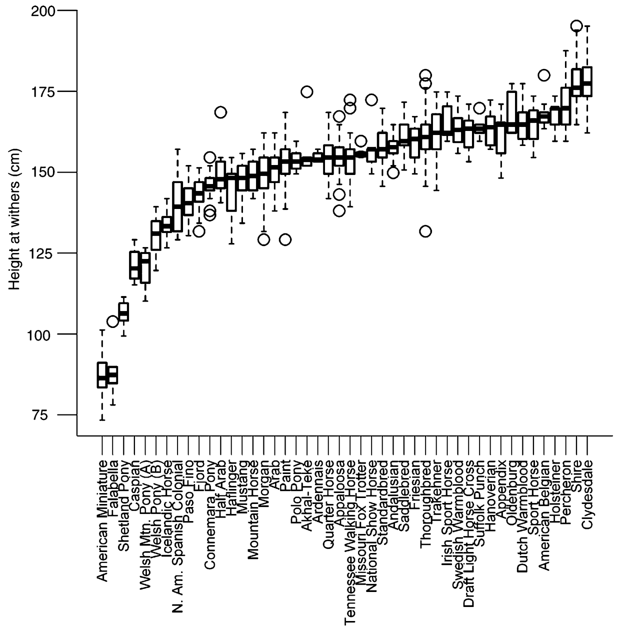

Figure 1 Median withers height for 46 breed populations in which a minimum of four horses were measured. Breeds are ranked left to right by median withers height. The box central bar indicates the median, the box ends delimit the 25th and 75th percentile, and the whiskers delimit the most extreme data point within 1.5 times the interquartile range.

where n is the number of body measurements we collected (Peres-Neto *et al.* 2005). With 33 measurements, this means our retained PCs should explain greater than ~3% of the variance. Principal component 1 (PC1) explains 65.9% of the variance and principal component 2 (PC2) explains 6.4%. Every one of the 33 body measurements has a positive factor-loading onto PC1, indicating a positive correlation (Fig. 2a). PC1 is therefore quantifying a proportional scaling of overall body size, where the head, neck, trunk and limbs increase or decrease in size coordinately.

Principal component 1 values for horses by breed are plotted in Fig. 3. The Falabella and American Miniature breeds have the smallest median PC1 values (Fig. 3, left side), while the Shire, Percheron, Clydesdale and American Belgian have the largest (Fig. 3, right side). PC1, by nature of its derivation from all 33 body measurements, quantifies a comprehensive assessment of body size information that no single measurement (like withers height, back length, heart girth or cannon length) is able to provide. Nevertheless, to the extent that single measurements are correlated with PC1 (Fig. 2a), they do provide some approximation of a horse's PC1 size score. However, there are key differences in the relative ranks of some breeds by withers height vs. by PC1. For example, while the Holsteiner is the 4th tallest breed at the withers, it ranks 8th in terms of overall size scored by PC1. Moreover, the Ardennais is the 6th largest breed by PC1 but only the 26th tallest breed by withers height.

The second PC explains a much smaller proportion of the variance than does PC1 (6.4% vs. 65.9%; see Fig. S4), indicating that it identifies a more subtle pattern of variation in body shape. PC2 nevertheless has a clear biological interpretation: it quantifies variation across the entire body in bone thickness and body shape. This is apparent from the pattern of factor loadings (Fig. 2b). A horse with a large positive PC2 score has a short and thick neck, short and thick limb bones, a relatively short back, a broad eye and muzzle width, and long hooves. In contrast, a horse with a negative PC2 score demonstrates the opposite pattern.

In addition to the set of quantitative body measures, we also collected factor scores for several traits, including the degree of feathering on the lower limbs and head profile (Fig. S2). Both of these traits are anecdotally correlated with

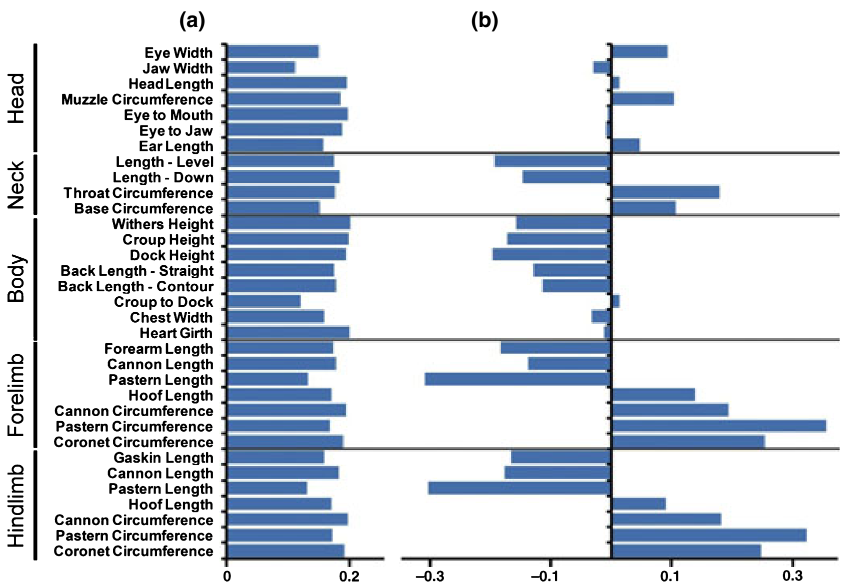

Figure 2 Principal components analysis factor loadings of body measurements onto principal component 1 (PC1) (a) and principal component 1 (PC2) (b) for the 1215 measured horses. PC1 factor loadings all have the same sign, while PC2 factor loadings for limb, back and neck length measures load oppositely (have different signs) from loadings for head, neck and limb circumference measures.

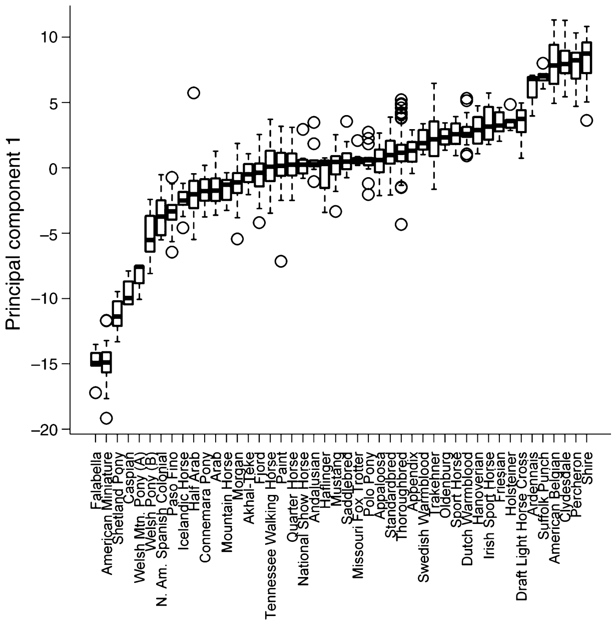

Figure 3 Box plots of each breed's distribution for the first principal component obtained from principal components analysis. The 46 breeds with a minimum of four measured horses are plotted. The box central bar indicates the median, the box ends delimit the 25th and 75th percentile, and the whiskers delimit the most extreme data point within 1.5 times the interquartile range.

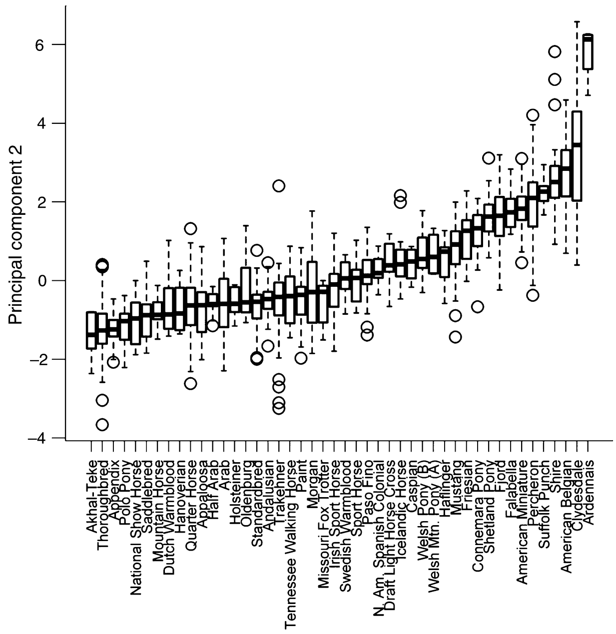

Figure 4 Box plots of each breed's distribution for the second principal component obtained from principal components analysis. The 46 breeds with a minimum of four measured horses are plotted. The box central bar indicates the median, the box ends delimit the 25th and 75th percentile, and the whiskers delimit the most extreme data point within 1.5 times the interquartile range.

body type in the horse; draft breeds often have 'Roman' nose profiles and copious feathering. Factor scores for both of these traits were associated with PC2 score in our dataset ($P < 0.0001$ for each, ANOVA).

We collected a subjective measure of bone thickness for each horse we measured and asked whether this could serve as a proxy for PC2. Observers could discern the difference between the very thickest horses and all others; the subjective factor score of '5', the thickest level, is predictive of a high PC2 score ($P < 0.0001$, ANOVA). However, observer scores for average or thin horses did not predict PC2 score well ($P = 0.6$, ANOVA), which illustrates the need for objective quantitative measurements to accurately capture this phenotype.

Discussion

Size and body conformation are critically important traits in nearly all horse breeds and are presumably under strong selection. Many breed registries select horses on functional criteria and encourage the breeding of horses with body types most suitable for those particular functions. The correct skeletal conformation is a key determinant of body type.

We find that body 'size' (PC1) and 'thickness' (PC2) have only limited variation within particular breeds compared to the total variation seen across our large panel of diverse breeds, consistent with selection within breed lines for preferred body types. The remaining variation within breeds is presumably because of ongoing segregation of causal alleles at contributory loci, measurement error and environmental factors.

We speculate that these data from 65 breeds may not have captured all skeletal variation in the horse species. Expansion of our sample sets from lightly sampled breeds and the addition of breeds that are geographically isolated, rare or under heavy selective pressure may provide additional patterns of skeletal variation.

PC1 is a single value for each horse that integrates measurement data from all over that horse's body; it can easily be applied to QTL studies. Furthermore, PC1 cleanly separates horse breeds based on the more traditional description of body type. Small ponies are grouped together with low PC1 scores (Fig. 3), light horses have mid-values and the large draft breeds all have high median PC1 scores. Unsurprisingly, the fine-boned Thoroughbred and Akhal-Teke breeds scored the lowest for PC2, while the heavy draft Ardennais and Clydesdale breeds have very high PC2 scores (Figs 4 & 5). Notably, however, many of the small pony breeds also have high PC2 scores, reflecting their relatively thick build for their small stature. Two of the smallest breeds sampled, the Falabella and the American Miniature, rank very close to the large draft breeds on the PC2 axis, despite their vast differences in PC1 size. With the aspect of 'body scale' removed by the first principal component, unexpected relationships like this can be clearly quantified so they can be put to use in genetic studies.

Figure 5 Two horses of similar height and body condition that exemplify the extremes of principal component 2 (PC2). Horse 'a', an Akhal-Teke, scored −1.73 for PC2, while horse 'b', an Ardennais, scored 6.04.

With genome-scale sequencing power increasing very rapidly, the 'limiting factor' in many genetic studies of the near future may no longer be genotyping but rather the acquisition of high-quality phenotypes. Further, the likelihood of mapping success will still be largely determined by the quality and characterization of samples and phenotypes (Carlson *et al.* 2004). Here, we have described skeletal variation in the horse species as two key axes of variation, PC1 (size) and PC2 (thickness), thus providing a solid foundation for future mapping studies that will define the genetic control of body size and shape in the horse.

Acknowledgements

Thanks to Bethany Austin, Debbie Cook, John Eberth, Caroline Salter, Kate McDermott, Karla Brown and Elizabeth Kowalski, as well as other members of the Brooks and Sutter Laboratories who volunteered their time and effort to measure horses. Our appreciation to the Equine Research Team at The Pennsylvania State University for their efforts in measuring horses. We thank JoAnn Bellone for providing photographs. We greatly appreciate funding for this project from Cornell University's Center for Vertebrate genomics, College of Agriculture and Life Sciences and the College of Veterinary Medicine. Finally, this project would not be possible without the support of horse owners who graciously provided their time and access to their animals.

Conflicts of interest

The authors have not declared any conflicts of interest.

References

Carlson C.S., Eberle M.A., Kruglyak L. & Nickerson D.A. (2004) Mapping complex disease loci in whole-genome association studies. *Nature* **429**, 446–52.

Dario C., Carnicella D., Dario M. & Bufano G. (2006) Morphological evolution and heritability estimates for some biometric traits in the Murgese horse breed. *Genetics and Molecular Research* **5**, 309–14.

Henneke D.R., Potter G.D., Kreider J.L. & Yeates B.F. (1983) Relationship between condition score, physical measurements and body fat percentage in mares. *Equine Veterinary Journal* **15**, 371–2.

Komosa M. & Purzyc H. (2009) Konik and Hucul horses: a comparative study of exterior measurements. *Journal of Animal Science* **87**, 2245–54.

Love S., Wyse C.A., Stirk A.J., Stear M.J., Calver P., Voute L.C. & Mellor D.J. (2006) Prevalence, heritability and significance of musculoskeletal conformational traits in Thoroughbred yearlings. *Equine Veterinary Journal* **38**, 597–603.

Peres-Neto P.R., Jackson D.A. & Somers K.M. (2005) How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics and Data Analysis* **49**, 974–97.

Sadek M.H., Al-Aboud A.Z. & Ashmawy A.A. (2006) Factor analysis of body measurements in Arabian horses. *Journal of Animal Breeding and Genetics* **123**, 369–77.

Sutter N.B., Bustamante C.D., Chase K. *et al.* (2007) A single IGF1 allele is a major determinant of small size in dogs. *Science* **316**, 112–5.

Weller R., Pfau T., May S.A. & Wilson A.M. (2006) Variation in conformation in a cohort of National Hunt racehorses. *Equine Veterinary Journal* **38**, 616–21.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Body measurements collected for each horse.

Figure S2 Histograms of counts of horses by (a) age; (b) body condition score where 1 = extremely underweight, 5 = ideal condition and 9 = overweight; (c) muzzle profile where 1 = deeply dish-faced (muzzle seen in profile curves in), 3 = straight and 5 = strongly Roman nosed (muzzle seen in profile curves out); (d) degree of feathering on lower limbs where 1 = no feathering and 5 = copious feathering; (e) bite conformation where 1 = extreme overbite and 5 = extreme underbite; and (f) subjective assessment of skeletal thickness without consideration of the horse's size, where 1 = a light, thin skeleton and 5 = a heavy, thick skeleton.

Figure S3 Coefficient of variation calculated for each measurement by breed. The box plot central bar indicates the median breed coefficient of variation.

Figure S4 Scree plot of the proportion of variance explained by each of the first ten principal components from the PCA performed on the 1215 horses measured.

Table S1 Count of measured horses by breed and sex.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.