

SCIENTIFIC REPORTS



OPEN

Spotted phenotypes in horses lost attractiveness in the Middle Ages

Saskia Wutke¹, Norbert Benecke², Edson Sandoval-Castellanos³, Hans-Jürgen Döhle⁴, Susanne Friederich⁴, Javier Gonzalez⁵, Jón Hallsteinn Hallsson⁶, Michael Hofreiter⁵, Lembi Lõugas⁷, Ola Magnell⁸, Arturo Morales-Muniz⁹, Ludovic Orlando¹⁰, Albína Hulda Pálsdóttir⁶, Monika Reissmann¹¹, Matej Ruttkay¹², Alexandra Trinks⁵ & Arne Ludwig¹

Received: 22 April 2016

Accepted: 09 November 2016

Published: 07 December 2016

Horses have been valued for their diversity of coat colour since prehistoric times; this is especially the case since their domestication in the Caspian steppe in ~3,500 BC. Although we can assume that human preferences were not constant, we have only anecdotal information about how domestic horses were influenced by humans. Our results from genotype analyses show a significant increase in spotted coats in early domestic horses (Copper Age to Iron Age). In contrast, medieval horses carried significantly fewer alleles for these phenotypes, whereas solid phenotypes (i.e., chestnut) became dominant. This shift may have been supported because of (i) pleiotropic disadvantages, (ii) a reduced need to separate domestic horses from their wild counterparts, (iii) a lower religious prestige, or (iv) novel developments in weaponry. These scenarios may have acted alone or in combination. However, the dominance of chestnut is a remarkable feature of the medieval horse population.

The variety of coat colour phenotypes is not only a valued feature of present-day domestic horses but has also fascinated humans ever since prehistoric times. Ancient cave paintings depicted several colour variants of wild horses, and further diversification occurred with domestication. In the early stages of horse domestication, a rapid increase in coat colour variation was enhanced by the artificial selection of horses under human care and their preferences for new phenotypes^{1,2}.

Today, several horse breeds are named after their coat colours and patterns, reflecting the great interest of humans in horse coloration. However, like most domestic animals, horses have undergone extensive breeding and breed standardization, especially during the last few centuries. Analysing only modern individuals can therefore lead to false conclusions about the history of breeds or the presence/absence of specific traits that may have existed in the past³, as has been recently demonstrated in studies on domestic chicken⁴ and horses⁵. Furthermore, human preferences have changed greatly over time and across cultures⁶. Based on our current knowledge, organized animal breeding over a larger region was first implemented in Europe at the time of the Roman Empire⁷. During the Roman period, long-distance exchange of individuals exhibiting favourable phenotypes was conducted to improve local stocks throughout the occupied territories^{7–9}. With the fall of the Roman Empire, animal husbandry in Europe became more disorganized and resulted in a wider phenotypic variation of domestic animals^{10,11}.

The relevance of coat coloration for investigating domestication has been previously demonstrated^{6,12–15}. For example, a strong effect of artificial selection on a coat colour locus was discovered in domestic pigs¹⁶. In horses, new phenotypes emerged soon after the onset of domestication¹². Subsequently, the alternating frequency of

¹Leibniz Institute for Zoo and Wildlife Research, Department of Evolutionary Genetics, 10315 Berlin, Germany.

²German Archaeological Institute, Department of Natural Sciences, Berlin, 14195 Berlin, Germany. ³Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Ciudad de México, Mexico. ⁴Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt – Landesmuseum für Vorgeschichte, 06114 Halle (Saale), Germany.

⁵University of Potsdam, Faculty of Mathematics and Natural Sciences, Institute for Biochemistry and Biology, 14476 Potsdam, Germany. ⁶The Agricultural University of Iceland, Faculty of Land and Animal Resources, IS-112 Reykjavik, Iceland. ⁷Archaeological Research Collection, Tallinn University, Rütli 10, 10130 Tallinn, Estonia. ⁸National Historical Museums, Contract Archaeology, 226 60 Lund, Sweden. ⁹Universidad Autónoma de Madrid, Laboratory of Archaeozoology, Madrid, Spain. ¹⁰Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350K Copenhagen, Denmark. ¹¹Humboldt University Berlin, Faculty of Life Sciences, Albrecht Daniel Thaer-Institute, 10115 Berlin, Germany. ¹²Slovak Academy of Sciences, Institute of Archaeology, 949 21 Nitra, Slovak Republic.

Correspondence and requests for materials should be addressed to A.L. (email: ludwig@izw-berlin.de)

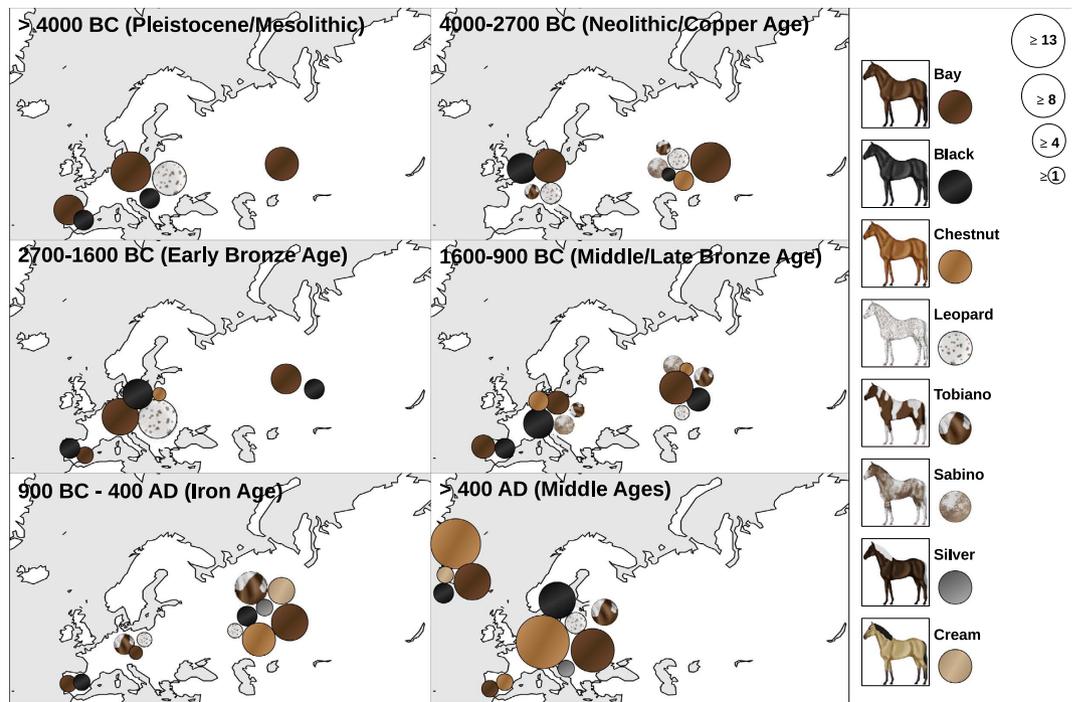


Figure 1. Occurrence of eight horse coat colour phenotypes based on their location for six different periods. For each spotted phenotype (leopard, tobiano and sabino) all samples, independent of their basic coat colour (bay, black or chestnut), are included. The map was made using Gimp 2.8.10 (www.gimp.org).

leopard complex spotting in ancient horses suggested changes in human preferences over time^{6,13}. However, because human preferences have varied over time and because horses started to be differentiated by use during the Middle Ages, a dense temporal and geographical sampling of horses will help establish a better understanding of the history of domestic horse populations. This study addresses the effect of human selection on coat colour phenotypes in horses from the Bronze Age to the Middle Ages, and therefore, has importance for our understanding of breeding preferences and the history of present breeds.

Results

In this study, we successfully genotyped 107 ancient samples for eight coat colour SNPs. Including data from previous studies^{12,13}, this dataset of 201 samples (Table S1) displays 14 different phenotypes (Table S4).

Figure 1 provides an overview of the number of different phenotypes for the respective region over six time periods. However, it should be noted that for spotted phenotypes (leopard, tobiano and sabino), all individuals are pooled irrespective of their basic coat colour (bay, black or chestnut). Early domestic horses (4000–2700 BC) showed six phenotypes, of which three were already present in pre-domestic horses (>4000 BC^{12,13}). In both the Bronze and Iron Ages (2700–900 BC and 900 BC–400 AD, respectively), nine phenotypes were already present (detailed information on phenotype numbers in Table S4).

Two Neolithic samples exhibited the tobiano phenotype, a colour pattern only occurring in domestic horses. One horse was from Salzmünde, Germany (cal. 3368–3101 BC), and the other from Botai, Kazakhstan (cal. 3654–3630 BC). Prior to this study, this phenotype had not been observed in horses before 1500 BC.

The differences in observed allele frequencies at different time periods were significant for different combinations of alleles and periods (see Supplementary Table S5). Alleles associated with new phenotypes, which became conspicuous in the early stages of domestication (chestnut ‘*e*’, tobiano ‘*KMI*’ and sabino ‘*SBI*’), decreased substantially in the Early Bronze Age. However, alleles for sabino and tobiano spotting (*SB1* and *KMI*) in particular increased again in frequency during the Late Bronze and Iron Ages, with *KMI* reaching an observed frequency of approximately 0.19 in the Iron Age (Figures S1–S3). Due to the occurrence of a few leopard alleles at the beginning of the medieval period, the inferred history of allele frequencies suggests that they were not decreasing during that time period. However, when the Middle Ages are separated into early and late periods, the change becomes significant, and the simulations show a strong decrease of spotted alleles driven by negative selection (Figure S5).

Most of the newly genotyped samples ($N = 107$) came from medieval European horses (after 400 AD, $N = 55$), which displayed a significant drop in the frequency of dilution or spotting phenotypes (Iron Age: 14 out of 31, Middle Ages: 8 out of 56, $p = 0.005$, based on Pearson’s chi-squared test) but a greater frequency of solid coat colours than horses of pre-Medieval Times (Fig. 2), which can also be observed in the lower allelic frequencies for most of the non-wild type alleles in the sample from the Middle Ages (Figure S1). Three medieval horses carried the pearl allele (*prl*). Since we did not detect this allele in any pre-medieval samples ($N = 52$), these samples

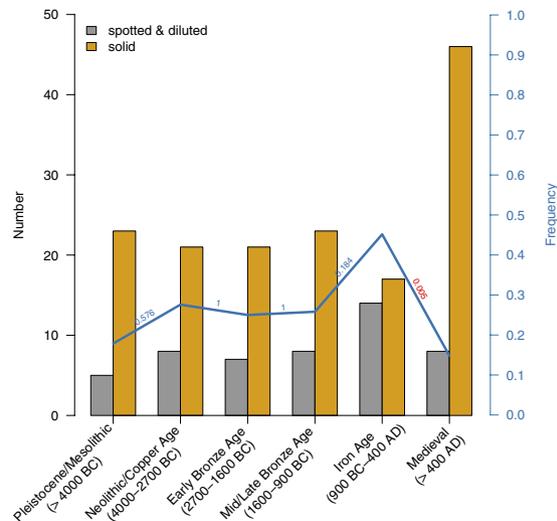


Figure 2. Numbers and frequencies of all spotted/diluted versus basic-coloured (solid) phenotypes for six different periods. The bar plot displays the absolute number of horses with the respective phenotype (grey represents all spotted and diluted horses, yellow represents horses with a solid coat colour). The blue line indicates the frequency of the spotted/diluted coat-colour variants, and the numbers above it indicate the significance of the changes in frequency, which were calculated using Pearson's chi-squared test.

currently represent the oldest detection of this allele. However, as heterozygous carriers, they did not exhibit the respective phenotype and, therefore, are not depicted in Fig. 1.

Although samples are neither chronologically nor geographically homogeneously distributed, there is no indication for any strong regional differences (Table S6). Medieval horses from various locations in Europe displayed similar frequencies of coat colour-associated alleles and their respective phenotypes. Nevertheless, Icelandic horses from the 9th–11th century (N = 21) did not harbour any spotted colour patterns despite this being a characteristic of current Icelandic horses.

In order to infer selection coefficients and temporal paths of allele frequencies we developed a novel MCMC method for time series of genetic data (see methods and Supplementary Appendix S1). It allowed inferring a striking rise in the frequency of the 'black' ASIP allele ('a') prior to domestication followed by an equally striking rise of the frequency for the chestnut allele. Figure 3 depicts the selection coefficients for the various phenotypes (basic coat colours in the upper row) as inferred by our method. With domestication, black and chestnut horses increased in frequency at the expense of bay horses, which is reflected in the significantly positive selection coefficients ($p < 0.01$). This trend was reversed in horses from the late Bronze and Iron Ages when the frequency and selection coefficient of chestnut horses increased, whereas the trend for black and bay phenotypes showed negative selection. Medieval horses again showed a significantly positive selection coefficient for black and chestnut alleles. This alternating pattern in selection coefficients is reflected in the respective allele frequencies for the derived ASIP and MC1R alleles ('a' and 'e', Figure S1), which have shown an apparent pattern of alternation since the onset of domestication (Figure S1).

For the dilution phenotypes, the information carried by the data (only 3–4 alleles were detected) was scarce providing limited information in the inference and showing a rise in Iron Age to Medieval times (Table S1; Figures S1 and S3).

Finally, after we incorporated the possibility that derived alleles (for novel phenotypes) "appear" spontaneously (emulating their new "appearance" by mutation) at some point between the beginning of the simulation and their first detection, we found that simulations in which the alleles newly emerged in the population, showed a much higher probability compared to those in which the allele was forced to exist since the beginning of the simulation. This finding strongly suggests that all derived alleles for diluted and spotted phenotypes appeared via mutation in the Holocene (after domestication), with the exception of the leopard complex spotting (*LP*) allele, which has existed since the Pleistocene (prior to domestication) supporting previous findings in horses and other domestic animals that the age of the alleles coding for colour variants only found in domestic animals tend to be very young when detected and likely occurred after domestication².

Discussion

Our extended dataset highlights important changes in human preferences for horse phenotypes. Specifically, the large number of samples from the Middle Ages demonstrate significant differences in coat colour phenotypes between pre-medieval and medieval horses. Spotted and diluted horses were considerably more frequent during the Bronze Age and Iron Age, whereas solid phenotypes, especially chestnut, were predominant in the Middle Ages. In addition, we discovered that tobiano spotting, which only occurs in domestic horses and had thus far only been detected after 1500 BC, was present in the Eneolithic/Copper Age (Kazakhstan, cal. 3654–3630 BC, and Germany, cal. 3368–3101 BC). Similar to chestnut and sabino spotting, the tobiano phenotype appears to arise

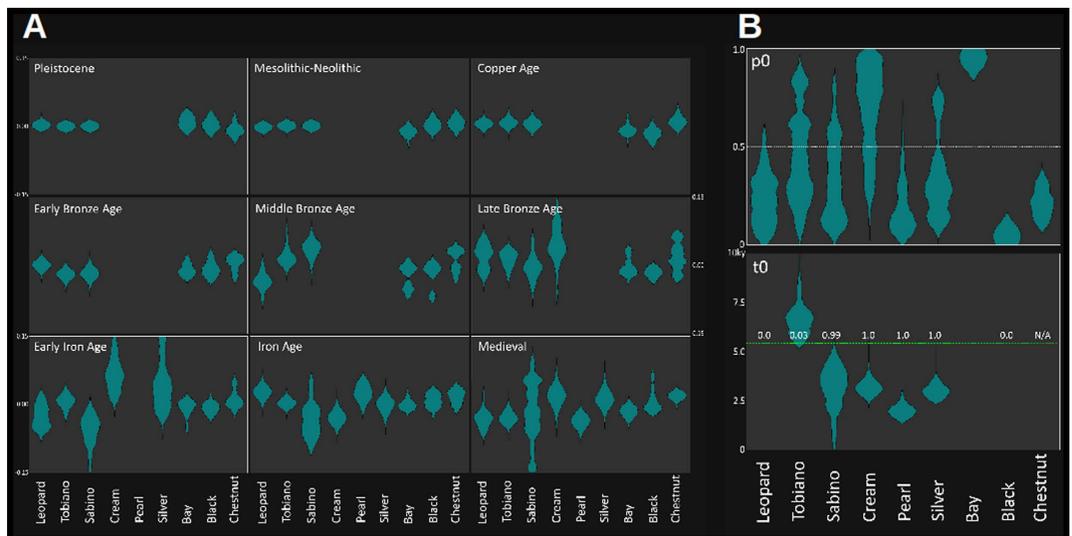


Figure 3. (A) Violin plots representing the posterior distributions of the parameters that were inferred by simulation-MCMC. Each chart represents a phenotype and each violin plot a selection coefficient in each period (indicated). The two parameters at right are the initial allele frequency and the age of the derived allele. The empty spaces (e.g. Cream phenotype in the Middle Bronze Age) occur because some simulations had several periods merged due to the derived allele appeared very late in the sampling, making the selection at the earliest periods pointless. (B) Violin plots of the initial allele frequency (top) and age (bottom) of the derived alleles. The empty places correspond to the cases when the derived allele first appeared in the Pleistocene, the allele associated to the corresponding phenotype was not derived (Bay); or the introduction of alleles didn't get an appreciable probability. The green line represents the time of domestication, and the numbers above the probabilities that the age of the derived allele was younger than domestication.

shortly after domestication, which is assumed to have started approximately 4000–3500 BC in the Ponto-Caspian steppe region (modern day Kazakhstan and Ukraine)¹⁷. That the tobiano allele most likely did not exist before domestication but emerged later is further supported by the estimated “time to the introduction” (age) of the allele. The same was observed for the Gait-keeper mutation in domestic horses, which has a rather recent origin in early medieval England¹⁸. Moreover, the detection of the tobiano phenotype in two Neolithic domestic horses from distant regions has important implications regarding the origins of horse domestication¹⁷ and their subsequent distribution. It also supports previous claims that the emergence of domestic horses in Central Europe at the end of the fourth millennium BC was facilitated by introduced horses from the Ponto-Caspian steppe¹⁹. The occurrence of the tobiano phenotype and its allelic frequency through time display a similar pattern as that of the leopard (*LP*) complex spotting allele⁶. In the Iron Age, the tobiano allele (*KM1*) occurred even more frequently than the *LP* allele. Both the high allele frequency and the fairly positive selection coefficient indicate that humans favoured spotted horses during the Iron Age. Still, as was the case for horses with the leopard spotting phenotype, tobiano horses were obviously not equally sought after during all time periods given their complete absence in the early and middle Bronze Age and their low prevalence in the early Middle Ages. Although the reason for these varying preferences through time cannot be resolved at this point, pleiotropic disadvantages may have played an important role^{1,20,21}.

In contrast to their Iron Age ancestors, early medieval horses displayed a strong reduction in spotted and diluted phenotypes. Such a decrease had apparently started in Roman times when, according to ancient Roman records, horses with uniform coat colour were preferred to spotted horses, as the latter were considered to be of inferior quality⁷. Historically, the early Middle Ages were marked by the collapse of the Roman Empire and large-scale human migrations across Eurasia and Northern Africa. The Roman Empire represented a cultural and economic entity characterized by, among other things, unprecedented agricultural advances. Its decline caused substantial changes in animal husbandry throughout the former empire, which was also observed in changes of livestock phenotypes^{10,11}. One might predict that less specialized breeding, with individuals not being as actively selected for desired traits and thus an increased effect of genetic drift, would further facilitate the reduced frequency of spotted coats and colour dilutions in horses. Considering possible pleiotropic disadvantages associated with some spotted and diluted phenotypes^{20–22}, especially in homozygous individuals^{22–24}, such a reduction might even have been enhanced by negative selection.

Spotted phenotypes might also have served as a visible cue to distinguish domestic from wild horses. With the decline of wild horse populations, the necessity to maintain a marker that set apart domestic animals, which had greater performance and higher economic value, might have been rendered obsolete.

Additionally, religious symbolism may have played a role in the decline of spotted horses during Medieval times. At the beginning of this period spotted horses appear to have been desirable. In the Apocalypse of St. John dated AD 81–96²⁵, where the end of the world revolves around the fall of Rome, the symbolism concerns the coat colour of the horse of each rider: the rider of victory has a white or white spotted horse, the rider of Famine a

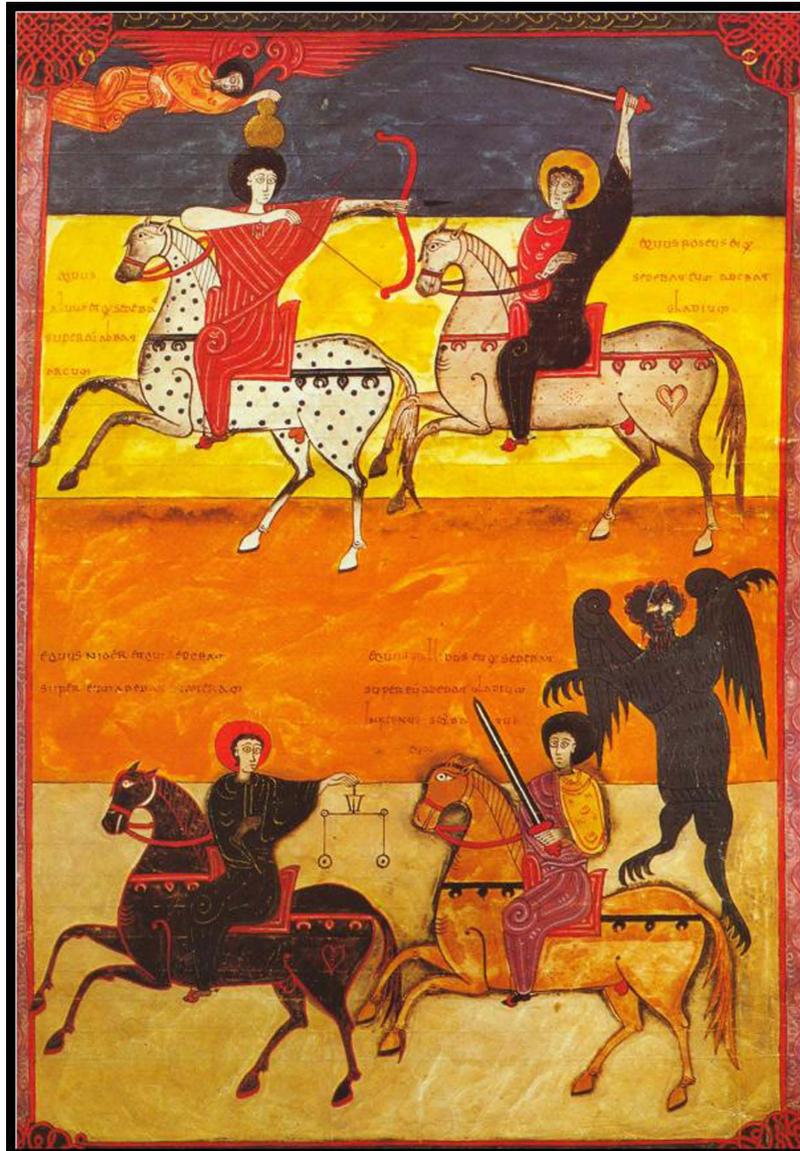


Figure 4. Four Horsemen of the Apocalypse on the *Beato de Fernando I y doña Sancha* dated 1047 AD. (Apoc. VI, 1–8f. 135; shelf 14-2 National Library, Madrid; https://commons.wikimedia.org/wiki/File:B_Facundus_135.jpg).

black horse, the rider of Death a bay horse and the rider of War a chestnut horse (Fig. 4). In other words, only the white/white spotted horse had a positive connotation, and this may explain the preferences of royalty for these horse phenotypes and the praise they received from authors such as Isidore of Seville^{26,27}. In time, the Church transmogrified the rider of victory into that of the Triumphant Church, a powerful icon of Medieval Christian symbolism²⁸. However, by an ironic twist of history, during the second half of the Middle Ages, after a series of epidemics -the Black Death among them- wiped out half of the European population, the single “good” rider was replaced by a fourth “bad” rider, that of the Plague²⁸, but the colours of the four horses remained the same resulting in the white or spotted horse now also being ridden by a rider with a negative connotation. Whether that switch meant that the negative connotation of the rider extended to the horse he was depicted on and spotted coats therefore became an undesirable “trait” for horses, which fostered the decline of spotted phenotypes remain open questions at this point, although the results of our study might grant support to this idea.

Notably, our study on ancient domestic horses contains a sampling bias because many medieval samples originate from graves of knights and noblemen. The high frequency of chestnut coat colour among these horses fits well with horse colours of the apocalyptic riders in which the rider of war had a chestnut horse. Moreover, chestnut horses are generally characterized as bold²⁹, which was definitely a beneficial behaviour for war horses. However, our samples do not necessarily represent working horses. Likely, the horses of the upper-class were bred for more fashionable traits. In the same manner as white horses were reserved for royalty² noblemen may have also had special preferences that are beyond our present-day understanding. Finally, improvements in weaponry

(i.e., the longbow) may have influenced the shift in coat colour preferences. A diluted or spotted horse may simply have been an easier target for enemies to see, especially over long distances or while moving.

We detected the allele for the pearl dilution (*prl*), an equine colour variant caused by a recessive mutation of the *MATP* gene, in an Iberian, a German and a Slovakian sample from the Middle Ages. In present-day horses, this phenotype occurs only in Iberian breeds and breeds that are descended from them (e.g., Peruvian Paso)³⁰. Although our results prove the existence of the pearl allele in horses outside the Iberian Peninsula in historic times, two factors argue for drift as a likely cause for its disappearance in non-Iberian horses: (i) the low observed frequency of this allele and (ii) the fact that the phenotype is visible only in homozygous individuals or in horses with an additional cream allele.

Strikingly, the Icelandic horses from the Viking Age show clear differences from their modern counterparts. The samples we analysed originated from graves around Iceland that date back to as early as the middle of the 9th century, shortly after Iceland was settled. These horses already carried the allele for Silver dapple (two of 19 individuals), a common trait in contemporary Icelandic horses, but they did not carry any alleles associated with spotted phenotypes, although such individuals are frequent in modern Icelandic horses^{31–33}. Therefore, we assume that spotted phenotypes were introduced after the 10th/11th century AD, which contradicts the popular claim that the import of horses to Iceland was prohibited for almost 1000 years³⁴.

In conclusion, the present study is the most comprehensive to date addressing coat colour differences in ancient horses. We found differential selection for spotted and solid phenotypes over time reflecting changes in human preferences. With the decline of the Roman Empire, solid phenotypes increased in frequency, whereas the frequency of spotted and diluted phenotypes decreased significantly. Although detailed inferences concerning geographic differences cannot be made, we observed considerable regional changes between medieval and present-day horses.

Methods

Ancient DNA analysis, authentication and sequencing. DNA extraction was performed in a specialized trace DNA laboratory at the Leibniz Institute for Zoo and Wildlife Research (Ludwig Laboratory) in Berlin, Germany, following standard procedures to avoid contamination. Independent replications were carried out in the ancient DNA laboratory at the University of Potsdam (Hofreiter Laboratory), Germany. To minimize contamination in all tissue, the surfaces of the samples were removed by abrasion. Next, 100–200 mg of bone or tooth was ground with a cryomill and incubated in 1 ml of an extraction buffer (0.45 M EDTA, 0.25 mg/ml proteinase K, pH 8.0) overnight with rotation at 37 °C. The following steps were conducted according to an optimized extraction protocol for short DNA fragments³⁵.

All samples were analysed for eight mutations located in six genes (*MC1R*, *ASIP*, *PMEL17*, *MATP*, *TRPM1*, *KIT*), including alleles for the basic coat colours (bay, black, chestnut), dilution phenotypes (cream, silver, pearl) and white spotting (tobiano, sabino, leopard). In addition to the ones described in previous studies^{12,13} a SNP in the *MATP* gene³⁰ associated with the Pearl phenotype was included (Table S2). Multiplex PCRs were performed in 20- μ l reaction volumes using a 4- μ l DNA extract and 1x AmpliTaq Gold PCR buffer II (ABI), 4 mM MgCl₂; 1 mg/ml bovine serum albumin (BSA); 250 μ M dATP, dCTP and dGTP; 500 μ M dUTP, 150 nM of each primer and 2 U of AmpliTaq Gold (ABI). To control for carryover contamination, each sample was treated with 1 U of heat-labile uracil-DNA glycosylase in an initial incubation step of 15 min at 37 °C. Target regions were amplified with target-specific primer pairs that had been tagged with common sequence tags (CS1 and CS2, Fluidigm).

Following multiplex amplification, PCR products were sequenced using a combination of multiplexed Illumina MiSeq sequencing and traditional Sanger capillary electrophoresis. For Illumina sequencing, the amplicons were indexed using a maximum of 384 barcodes (containing Illumina adaptor sequences) in an additional amplification step and then pooled. To remove primer and adaptor dimers and other amplification products, the multiplex PCR products were purified using the AMPure PCR purification system (Agencourt) with a 1.8-fold ratio of SPRI beads relative to the reaction volume. The pooled and purified amplicon libraries were then sequenced on an Illumina MiSeq sequencer using CS1 and CS2 as well as their reverse complements as sequencing primers according to the manufacturer's instructions. For a few samples, the PCR products were generated with the same target-specific primer pairs and then sequenced on an ABI 3130xl Genetic Analyser using the BigDye[®] Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). In addition to the sequencing analyses, the 11 bp deletion in the *ASIP* gene was genotyped via gel electrophoresis.

Data Analysis. Traditional Sanger sequences were edited and aligned in Geneious v8.1 (Biomatters). Illumina reads were automatically demultiplexed by the Illumina MiSeq software and further processed by applying the following steps: Illumina sequencing adaptors were cut from the reads using cutadapt³⁵; reads were then filtered for quality with Trimmomatic 0.33^{36,37}; all reads that contained more than 8 bp with a Phred score below 20 were discarded and only reads with a minimum length of 50 bp were retained. The trimmed and filtered paired-end reads were merged into a single read with FLASh³⁸ to improve their overall quality. Using Bowtie 2³⁹ and samtools 0.1.19⁴⁰, the merged reads were mapped to the reference sequences (accession numbers in Table S2) for the respective PCR products. For each sample and locus, a consensus sequence was generated in Geneious v8.1 (Biomatters) from which the genotype and the resulting coat colour phenotype was deduced. To investigate the changes of the number of spotted and solid phenotypes between the different times Pearson's chi-squared test was applied.

Allelic dropout. The probability (P) of a false homozygous individual per locus was calculated after n replicates as follows: $P = K(K/2)^{n-1}$, where K is the observed number of allelic dropouts/false homozygote divided by all heterozygous individuals⁴¹. We did a minimum of three replications, which reduced the risk of the non-detection of a heterozygous individual to 0.2%.

Temporal test. We applied a Bayesian test of temporal changes in allele frequencies⁴² in order to evaluate the null hypothesis that the observed changes in allele frequencies could be attributed solely to genetic drift and sampling error. For doing so, it was necessary to group the individuals in samples with a common (average) age (see Figure S2). Despite some bias is introduced due to the averaging of ages when grouped, this test can provide clues regarding the potential presence of selection.

Selection coefficients. To determine how natural and artificial selection could have changed over time, we divided the sampling timeframe in nine periods (Pleistocene, Mesolithic/Neolithic, Copper Age, Early Bronze Age, Middle Bronze Age, Late Bronze Age, Early Iron Age, Iron Age and Medieval) based on historical criteria; but fusing some periods when sample sizes demanded it. Then, we estimated a selection coefficient for each phenotype and period.

The large number of unknown parameters, together with the epistatic and dependency relationships between genes and phenotypes (Figure S4) precluded the use of previously reported methods for detecting selection (e.g. ref. 43), as well as the use of approximate Bayesian computation (ABC), which failed to provide informative posteriors even after we implemented a Sequential Monte Carlo (SMC) procedure that theoretically improved the efficiency of by a factor of 10^9 .

For that reason, we implemented a novel MCMC algorithm which consisted in the simulation of the temporal path of population allele frequencies under a Wright-Fisher model, followed by the analytic calculation of the likelihood of the empirical sample (which employed the simulated allele frequencies). Then we implemented an inference by MCMC based on the likelihoods of the samples given the simulated population frequencies (see Appendix S1 for details). As a Bayesian Monte Carlo method, our method allows to account for relevant sources of uncertainty in the observed data and other variables. This allowed treating properly the uncertainty associated to the effective population sizes, generation time and the ages of the samples, which uncertainty came from stratigraphic or radiocarbon dating.

The MCMC was run under a Metropolis-Hasting algorithm with a Gibbs sampler that was implemented to optimize the exploration of the large parametric space of selection coefficients. The phenotypes or genes that had epistatic or dependency relationships were run together; otherwise the genes were run independently. For each inference we ran 20 chains, 2.0×10^4 steps each (2.0×10^5 for ASIP-MC1R), with a burnin of 50%. Specifics about the MCMC are shown in Tables S7. Simulations were programmed following previously published methods^{6,44}, and modifications were programmed to account for the genes and phenotypes interactions. Following literature, the simulations incorporated a growth in horse population, which began with the onset of domestication^{1,45–50}. The priors for population sizes, generation time and time to the start of the population growth are shown in Table S7. Ages were sampled from normal or uniform priors to take uncertainties in the radiocarbon or stratigraphic age into account, and four samples from the Late Pleistocene were sampled from exponential priors because of their very large stratigraphic range (11.7–126 ka); and the fact that their genotyping was successfully (because ancient DNA decline is exponential) means that their ages are more likely younger than older.

The simulations implemented two models to deal with the initial state of derived alleles, allowing them to exist since the beginning of the simulated time frame (at a frequency being an estimating parameter) or being introduced by mutation at some point between the beginning of the simulated time frame and its first detection (its age being an estimating parameter). The selection of the right model was performed implicitly by allowing the MCMC to shift among models (see ref. 51 and Appendix S1).

The simulations were programmed using the code of the programs employed in refs 6 and 41 as template. The source code is available from the following link: <https://www.dropbox.com/sh/uns1gsprj7xbbzo/AAAEr4hfya16P5nejpXVVnUfa?dl=0>.

References

- Cieslak, M., Reissmann, M., Hofreiter, M. & Ludwig, A. Colours of domestication. *Biol. Rev.* **86**, 885–899 (2011).
- Linderholm, A. & Larson, G. The role of humans in facilitating and sustaining coat colour variation in domestic animals. *Semin. Cell Dev. Biol.* **24**, 587–593 (2013).
- Hofreiter, M. *et al.* The future of ancient DNA: Technical advances and conceptual shifts. *BioEssays* **37**, 284–93 (2015).
- Girdland Flink, L. *et al.* Establishing the validity of domestication genes using DNA from ancient chickens. *Proc. Natl. Acad. Sci.* **111**, 6184–6189 (2014).
- Der Sarkissian, C. *et al.* Evolutionary Genomics and Conservation of the Endangered Przewalski's Horse. *Curr. Biol.* **25**, 2577–2583 (2015).
- Ludwig, A. *et al.* Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses. *Philos. Trans. R. Soc. London B* **370**, (2015).
- Peters, J. *Römische Tierhaltung und Tierzucht: eine Synthese aus archäozoologischer Untersuchung und schriftlich-bildlicher Überlieferung* 5 (Leidorf, 1998).
- Albarella, U., Johnstone, C. & Vickers, K. The development of animal husbandry from the Late Iron Age to the end of the Roman period: a case study from South-East Britain. *J. Archaeol. Sci.* **35**, 1828–1848 (2008).
- Ludwig, A. *et al.* Tracing the genetic roots of the indigenous White Park Cattle. *Anim. Genet.* **44**, 383–386 (2013).
- Schibler, J. & Schlumbaum, A. History and economic importance of cattle (*Bos taurus*, L.) in Switzerland from Neolithic to Early Middle Ages. *Schweiz. Arch. Tierheilkd.* **149**, 23–29 (2007).
- Crabtree, P. J. Animal husbandry and farming in East Anglia from the 5th to the 10th centuries CE. *Quat. Int.* **346**, 102–108 (2014).
- Ludwig, A. *et al.* Coat Color Variation at the Beginning of Horse Domestication. *Science (80-)*. **324**, 485 (2009).
- Pruvost, M. *et al.* Genotypes of predomestic horses match phenotypes painted in Paleolithic works of cave art. *Proc. Natl. Acad. Sci.* **108**, 18626–18630 (2011).
- Svensson, E. M. *et al.* Tracing genetic change over time using nuclear SNPs in ancient and modern cattle. *Anim. Genet.* **38**, 378–383 (2007).
- Ollivier, M. *et al.* Evidence of coat color variation sheds new light on ancient canids. *PLoS One* **8**, e75110 (2013).
- Fang, M., Larson, G., Soares Ribeiro, H., Li, N. & Andersson, L. Contrasting Mode of Evolution at a Coat Color Locus in Wild and Domestic Pigs. *PLoS Genet.* **5**, e1000341 (2009).
- Outram, A. K. *et al.* The earliest horse harnessing and milking. *Science (80-)*. **323**, 1332–5 (2009).

18. Wutke, S. *et al.* The origin of ambling horses. *Curr. Biol.* **26**, R697–R699 (2016).
19. Benecke, N. Zu den Anfängen der Pferdehaltung in Eurasien. Aktuelle archäozoologische Beiträge aus drei Regionen. *Ethnogr. Zeitschrift* **43**, 187–226 (2002).
20. Bellone, R. R. Pleiotropic effects of pigmentation genes in horses. *Anim. Genet.* **41**, 100–110 (2010).
21. Bellone, R. R. *et al.* Evidence for a Retroviral Insertion in TRPM1 as the Cause of Congenital Stationary Night Blindness and Leopard Complex Spotting in the Horse. *PLoS One* **8**, e78280 (2013).
22. Reissmann, M. & Ludwig, A. Pleiotropic effects of coat colour-associated mutations in humans, mice and other mammals. *Semin. Cell Dev. Biol.* **24**, 576–586 (2013).
23. Hedrick, P. W. What is the evidence for heterozygote advantage selection? *Trends Ecol. Evol.* **27**, 698–704 (2012).
24. Hedrick, P. W. Heterozygote Advantage: The Effect of Artificial Selection in Livestock and Pets. *J. Hered.* **106**, 141–154 (2015).
25. Stuckenbruck, L. T. *Revelation in Commentary on the Bible* (ed. Dunn, J. D. G., Rogerson, J. W.) 1535–1536 (Eerdmans 2003).
26. Barney, S. A., Lewis, W. J., Beach, J. A. & Berghof, O. *The etymologies of Isidore of Seville*. (Cambridge University Press, 2006).
27. Echegaray, J. G., del Campo Hernández, A., Freeman, L. G. & Soto, J. L. C. *Obras completas y complementarias de Beato de Liébana. I. Comentario al Apocalipsis. Himno O Dei Verbum. Apologético: I. 1* (Biblioteca de autores cristianos, 2004).
28. Clark, W. B. *A medieval book of beasts: the second-family bestiary: commentary, art, text and translation*. (Woodbridge, 2013).
29. Finn, J. L. *et al.* The relationship between coat colour phenotype and equine behaviour: A pilot study. *Appl. Anim. Behav. Sci.* **174**, 66–69 (2016).
30. Sponenberg, D. P. *Equine Color Genetics* (Wiley, 2009).
31. Brooks, S. A., Lear, T. L., Adelson, D. L. & Bailey, E. A chromosome inversion near the KIT gene and the Tobiano spotting pattern in horses. *Cytogenet. Genome Res.* **119**, 225–230 (2007).
32. Hauswirth, R. *et al.* Mutations in MITF and PAX3 cause 'splashed white' and other white spotting phenotypes in horses. *PLoS Genet.* **8**, e1002653 (2012).
33. Haase, B. *et al.* Seven novel KIT mutations in horses with white coat colour phenotypes. *Anim. Genet.* **40**, 623–629 (2009).
34. Nolf, P. S. Detecting Icelandic Horse Origins. *Icelandic Horse Q.* 18–23 (2012).
35. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci.* **10**, 15758–15763 (2013).
36. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
37. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, doi: 10.1093/bioinformatics/btu170 (2014).
38. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
39. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
40. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
41. Gagneux, P., Boesch, C. & Woodruff, D. S. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Mol. Ecol.* **6**, 861–868 (1997).
42. Sandoval-Castellanos, E. Testing temporal changes in allele frequencies: a simulation approach. *Genet. Res. (Camb.)* **92**, 309–320 (2010).
43. Schraiber, J. G., Evans, S. N. & Slatkin, M. Bayesian inference of natural selection from allele frequency time series. *Genetics* **203**, 493–511 (2016).
44. Nielsen, R. & Wakeley, J. Distinguishing Migration From Isolation: A Markov Chain Monte Carlo Approach. *Genetics* **158**, 885–896 (2001).
45. Gerbault, P. *et al.* In *Population Dynamics in Pre- and Early History New Approaches by Stable Isotopes and Genetics* (eds Kaiser, E., Burger, J. & Schier, W. O.) 17–30 (De Gruyter, 2012).
46. Lippold, S., Matzke, N. J., Reissmann, M. & Hofreiter, M. Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evol. Biol.* **11**, 328 (2011).
47. Achilli, A. *et al.* Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proc. Natl. Acad. Sci.* **109**, 2449–2454 (2012).
48. Warmuth, V. *et al.* Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proc. Natl. Acad. Sci.* **109**, 8202–8206 (2012).
49. Lippold, S. *et al.* Discovery of lost diversity of paternal horse lineages using ancient DNA. *Nat. Commun.* **2**, 450 (2011).
50. Librado, P. *et al.* Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl. Acad. Sci.* **112**, E6889–E6897 (2015).
51. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (LU 852/7–4). We thank Pavel Kosintsev for ancient horse samples from the Caspian Sea region. We thank Dietmar Lieckfeldt and Melanie Pruvost for their assistance as well as Dorina Lenz for helping with the data analyses. We also acknowledge the valuable comments from Dolores Carmen Morales-Muñiz about the Riders of the Apocalypse.

Author Contributions

A.L., M.H., N.B. and M.Re. planned the project. A.L. designed the study. A.L., M.Re. and S.W. designed the experiments. S.W. performed the experiments. J.G. and A.T. performed replications. H.-J.D., S.F., J.H.H., P.K., L.L., O.M., A.M.-M., L.O., A.H.P. and M. Ru. provided samples and information about the archaeological/historical context of the samples. S.W. and E.S.-C. analysed the data. E.S.-C. modelled the selection coefficients. A.L., S.W. and E.S.-C. wrote the draft. All authors contributed to writing the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wutke, S. *et al.* Spotted phenotypes in horses lost attractiveness in the Middle Ages. *Sci. Rep.* **6**, 38548; doi: 10.1038/srep38548 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016