



Worldwide frequency distribution of the 'Gait keeper' mutation in the *DMRT3* gene

M. Promerová*¹, L. S. Andersson^{†1}, R. Juras[‡], M. C. T. Penedo[§], M. Reissmann[¶], T. Tozaki^{**}, R. Bellone^{††}, S. Dunner^{‡‡}, P. Hořín^{§§}, F. Imsland*^{¶¶}, P. Imsland^{¶¶¶}, S. Mikko[†], D. Modrý^{***†††}, K. H. Roed^{‡‡‡}, D. Schwochow[†], J. L. Vega-Pla^{§§§}, H. Mehrabani-Yeganeh^{¶¶¶¶}, N. Yousefi-Mashouf^{¶¶¶¶}, E. G. Cothran[‡], G. Lindgren[†] and L. Andersson*[†]

*Department of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden. [†]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden. [‡]Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843-4458, USA. [§]Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, CA 95616, USA. [¶]Breeding Biology and Molecular Genetics, Faculty of Agriculture and Horticulture, Humboldt University, 10115 Berlin, Germany. ^{**}Genetic Analysis Department, Laboratory of Racing Chemistry, Utsunomiya, Tochigi 320-0851, Japan. ^{††}Department of Biology, University of Tampa, Tampa 33606, USA. ^{‡‡}Veterinary Faculty, Universidad Complutense de Madrid, 28040 Madrid, Spain. ^{§§}Institute of Animal Genetics, CEITEC, University of Veterinary and Pharmaceutical sciences, 612 42 Brno, Czech Republic. ^{¶¶}Menntaskolinn vid Hamrahlid, Hamrahlid 10, 105 Reykjavík, Iceland. ^{***}Department of Pathology and Parasitology, Faculty of Veterinary Medicine, CEITEC, University of Veterinary and Pharmaceutical sciences, 612 42 Brno, Czech Republic. ^{†††}Institute of Parasitology, Biology Centre of Academy of Sciences of the Czech Republic, v.v.i., 612 42 Česká Budějovice, Czech Republic. ^{‡‡‡}Department of Basic Sciences & Aquatic Medicine, Norwegian School of Veterinary Science, N-0033 Oslo, Norway. ^{§§§}Laboratorio de Investigación Aplicada, Cría Caballar de las Fuerzas Armadas, 14080 Cordoba, Spain. ^{¶¶¶¶}Department of Animal Science, University of Tehran, 54500 Tehran, Iran.

Summary

For centuries, domestic horses have represented an important means of transport and served as working and companion animals. Although their role in transportation is less important today, many horse breeds are still subject to intense selection based on their pattern of locomotion. A striking example of such a selected trait is the ability of a horse to perform additional gaits other than the common walk, trot and gallop. Those could be four-beat ambling gaits, which are particularly smooth and comfortable for the rider, or pace, used mainly in racing. Gaited horse breeds occur around the globe, suggesting that gaitedness is an old trait, selected for in many breeds. A recent study discovered that a nonsense mutation in *DMRT3* has a major impact on gaitedness in horses and is present at a high frequency in gaited breeds and in horses bred for harness racing. Here, we report a study of the worldwide distribution of this mutation. We genotyped 4396 horses representing 141 horse breeds for the *DMRT3* stop mutation. More than half (2749) of these horses also were genotyped for a SNP situated 32 kb upstream of the *DMRT3* nonsense mutation because these two SNPs are in very strong linkage disequilibrium. We show that the *DMRT3* mutation is present in 68 of the 141 genotyped horse breeds at a frequency ranging from 1% to 100%. We also show that the mutation is not limited to a geographical area, but is found worldwide. The breeds with a high frequency of the stop mutation (>50%) are either classified as gaited or bred for harness racing.

Keywords ambling, domestication, horse, locomotion, pace, running walk

Address for correspondence

L. Andersson, Department of Medical Biochemistry and Microbiology, Uppsala University, Box 582, SE-75123 Uppsala, Sweden.
E-mail: leif.andersson@imbim.uu.se

¹Equal contribution.

Accepted for publication 19 November 2013

Introduction

Few animals have been of such great value to humans as horses when it comes to means of transportation. All over the world, horses have been used for everyday transportation, in military settings, cattle herding and agricultural power, pulling carriages and carts, pleasure riding or racing. Over the centuries, horse populations and breeds

have been shaped by humans based on the different purposes for which the animals were used. For instance, in breeds used for riding over long distances, as those used by nomads and cattle herders, it was important that the horse moved in a way comfortable for the rider. Later, these traits also were valued for trail riding as well as show performances. In horses used for harness racing, breeders selected for the ability of a horse to trot or pace at high speed.

The ambling gaits are particularly comfortable for the rider, and horse breeds exhibiting these gaits are referred to as gaited. Some gaited horses also are able to perform pacing, which is a lateral two-beat gait. The ambling gaits have a four-beat rhythm with the footfall pattern of a walk, but often reaching speeds comparable to, or even exceeding, the trot without any aerial phase typical of the faster gaits (Robilliard *et al.* 2007). Ambling gaits are classified into numerous types, nearly as many as there are gaited breeds. Gaits are often characteristic of a particular breed (for a kinematic study see Nicodemus & Clayton 2003). Some horses are not able to learn a desired gait, others require extensive training, whereas some have a natural talent for ambling gaits. In some breeds, mostly from the New World, such as the Puerto Rican Paso Fino or the Tennessee Walking Horse, very strong selection has been applied, and basically all individuals of the breed are naturally gaited. In other breeds, such as the American Saddlebred or Mangalarga Marchador, both types of performances are utilised; thus, a fraction of the population is gaited and the rest cannot perform the ambling gait(s). By contrast, in many breeds, such as horses bred for show jumping or high-speed gallop racing, a pacing or ambling phenotype is considered inferior, and these traits have been strongly selected against.

Until recently, the genetic basis and the mode of inheritance of gaitedness were unknown. However, one of the founders of genetics, William Bateson, proposed as early as 1907 based on breeding records that pacing may be a recessive trait in horses (Bateson 1907). A recent study in fact demonstrated that a single-base substitution in the *double-sex and mab-3-related transcription factor 3* (*DMRT3*) gene has a major impact on the ability of a horse to pace and amble (Andersson *et al.* 2012), revealing the strong genetic basis for this trait. The mutation (cytosine to adenine), initially discovered in Icelandic horses, causes a premature stop codon and thus a truncation of the *DMRT3* protein (Andersson *et al.* 2012). The mutation (*DMRT3_Ser301STOP*) also was denoted *Gait keeper* due to the strong effect on the pattern of locomotion. *DMRT3* is one of the eight known genes that are homologous to the *doublesex* gene of *Drosophila* and the *mab-3* gene of *Caenorhabditis elegans* (Raymond *et al.* 1998). These genes encode transcription factors with a DM DNA-binding domain. The most well-studied members of this gene family have an established role in control of sex differentiation, but the expression of some members of the family in tissues other than

gonads suggests other functions (Hong *et al.* 2007). In fact, *DMRT3* is expressed in neuronal cells, and detailed studies in mice showed that it is expressed in inhibitory interneurons with projecting ipsi- and contralateral axons that make direct synaptic connections to motor neurons present in a specific region of the spinal cord (Andersson *et al.* 2012). Furthermore, the characterisation of *Dmrt3* knockout mice confirmed that the *DMRT3* protein plays a pivotal role for coordinating limb movements (Andersson *et al.* 2012).

Screening several horse breeds for the *DMRT3* mutation revealed that the occurrence of this mutation is dichotomous with an allele frequency of nearly 100% in six tested gaited breeds and zero in seven tested non-gaited breeds and in the Przewalski's horse (Andersson *et al.* 2012). In addition to the strong association with ambling gaits, this mutation was shown to occur in breeds used for harness racing, pacers as well as trotters. The mutated allele was shown to be associated with superior racing performance in Swedish Standardbred trotters and to be fixed in American Standardbred trotters and pacers (Andersson *et al.* 2012). This finding suggests a more complex effect of the *DMRT3* nonsense mutation on locomotion in horses by extending the range of speed at which horses keep a symmetrical gait, such as trot or pace, instead of switching to the asymmetric gallop, the ancestrally preferred gait at high speeds.

In this study, we genotyped the *DMRT3* stop mutation in additional horse breeds with wide geographical distribution. Our goal was (i) to investigate in which parts of the world the mutation occurs, (ii) to investigate whether the distribution of the mutation follows the distribution pattern of gaited horses in a wider sample set than previously studied and (iii) to test whether the mutation is strictly restricted to gaited breeds or might occur also in breeds considered non-gaited.

Material and methods

Sample collection

We obtained (i) extracted DNA from the Animal Genetics Laboratory, SLU, Sweden, and (ii) DNA, hair or blood samples collected by collaborators, provided by horse owners or archived at the Veterinary Genetics Laboratory, University of California, Davis, USA. DNA from blood samples was extracted using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instruction. DNA from hair samples was extracted using standard hair-preparation methods. Details of breeds and sample sizes are provided in Table 1. A total number of 4396 horses representing 141 horse breeds were included in the study.

SNP genotyping, PCR and Sanger sequencing of the *DMRT3* coding region

The *DMRT3_Ser301STOP* mutation (chr23: 22 999 655 bp) and a SNP 32 kb upstream of it (*BIEC2_620109*, chr23:

Table 1 Genotype and allele frequencies of the *DMRT3_Ser301STOP* mutation in different horse breeds. Breeds are sorted based on gaitedness (where this information was available) and alphabetical order of the breed names.

Breeds	Country of origin ¹	n	Genotype			Frequency of A allele (%)	Chi-square values for HWE test ²	Gaitedness ³
			AA	CA	CC			
American Paso Fino	USA	34	31	3	0	95.6	0.07	Gaited – paso
Campolina	Brazil	18	15	2	1	88.9	3.45	Gaited – marcha
Colombian Paso Fino	Colombia	80	75	1	4	94.4	62.27***	Gaited – paso
Cretan Horse	Crete, Greece	66	63	3	0	97.7	0.04	Gaited
Florida cracker	USA	24	15	2	7	66.7	15.84***	Gaited – singlefoot
Hokkaido Horse	Japan	48	28	13	7	71.9	5.23*	Gaited
Icelandic Horse ⁴	Iceland	219	117	93	9	74.7	3.27	Gaited – tölt, pace
Kentucky Mountain Saddle Horse ⁴	USA	25	21	4	0	92.0	0.19	Gaited – singlefoot
Missouri Fox Trotter ⁴	USA	42	42	0	0	100.0	NA	Gaited – foxtrot, pace
Peneia (Pinia)	Greece	17	16	1	0	97.1	0.02	Gaited
Peruvian Paso ⁴	Peru	22	22	0	0	100.0	NA	Gaited – paso
Puerto Rican Paso Fino ⁴	Puerto Rico	78	77	1	0	99.4	0.00	Gaited – paso
Rocky Mountain Horse ⁴	USA	27	27	0	0	100.0	NA	Gaited – singlefoot, rack
Tennessee Walking Horse ⁴	USA	54	54	0	0	100.0	NA	Gaited – running walk
American Curly	USA	45	0	15	30	16.7	1.80	Some – foxtrot
American Saddlebred	USA	89	6	37	46	27.5	0.16	Some – rack, slow gait
Appaloosa	USA	20	1	3	16	12.5	1.98	Some
Asturcon	Spain	24	0	0	24	0.0	NA	Some?
Basuto Pony	South Africa	30	0	4	26	6.7	0.15	Some?
Boer Pony	South Africa	20	0	6	14	15.0	0.62	Some
Brazilian Criollo	Brazil	21	0	1	20	2.4	0.01	Some
Colombian Trocha Pura	Colombia	67	2	10	55	10.4	2.74	Some – trocha
Colombian Criollo Trocha y Galope	Colombia	4	0	2	2	25.0	NA	Some – trocha
Faeroe Pony	Faeroe Islands	21	0	6	15	14.3	0.58	Some – tölt, pace
Kirgiz	Kyrgyzstan	31	7	11	13	40.3	2.14	Some – pace
Lewitzer	Germany	20	0	3	17	7.5	0.13	Some – pace
Mangalarga	Brazil	14	0	2	12	7.1	0.08	Some – marcha
Mangalarga Marchador	Brazil	22	5	10	7	45.5	0.15	Some – marcha
Marajo	Brazil	14	0	2	12	7.1	0.08	Some
Marwari	India	9	0	0	9	0.0	NA	Some – revaal
Miniature Horse	USA	109	1	4	104	2.8	10.78**	Some – pace
Mongolian	Mongolia	134	2	8	124	4.5	12.23***	Some
Morgan	USA	50	1	12	37	14.0	0.00	Some – singlefoot
Newfoundland	Canada	26	0	6	20	11.5	0.44	Some
Nooitgedacht	South Africa	14	1	4	9	21.4	0.32	Some
Pindos	Greece	15	1	5	9	23.3	0.07	Some
Pura Raza Galega	Spain	3	1	1	1	50.0	NA	Some
Rhodes	Greece	6	1	3	2	41.7	NA	Some
Shackleford Banks	USA	41	1	7	33	11.0	0.65	Some
Sierra Tarahumara	Mexico	18	2	3	13	19.4	3.94*	Some
Spanish Mustang	USA	15	0	3	12	10.0	0.19	Some
Spanish Pure Breed	Spain	46	0	1	45	1.1	0.01	Some?
Turkmen	Iran	20	0	4	16	10.0	0.25	Some – pace
Welsh Cob	Great Britain	11	0	0	11	0.0	NA	Some – pace
Venezuelan Criollo	Venezuela	21	0	7	14	16.7	0.84	Some
Akhal-Teke	Turkmenistan	43	0	0	43	0.0	NA	Not gaited
Altai	Russia (West Siberia)	25	0	4	21	8.0	0.19	Not gaited
Arabian ⁴	Middle East, Syria, Iran	69	0	0	69	0.0	NA	Not gaited
Ardennes ⁴	Belgium	39	0	0	39	0.0	NA	Not gaited
Barb	North Africa	15	0	0	15	0.0	NA	Not gaited
Black Forest Chestnut	Germany	15	0	0	15	0.0	NA	Not gaited
Breton	France	15	0	0	15	0.0	NA	Not gaited
British Riding Pony	Great Britain	1	0	0	1	0.0	NA	Not gaited
Camargue	France	3	0	0	3	0.0	NA	Not gaited
Canadian	Canada	14	0	0	14	0.0	NA	Not gaited

Table 1 (continued)

Breeds	Country of origin ¹	n	Genotype			Frequency of A allele (%)	Chi-square values for HWE test ²	Gaitedness ³
			AA	CA	CC			
Caspian	Iran	52	1	7	44	8.7	1.15	Not gaited
Clydesdale	Great Britain	17	0	0	17	0.0	NA	Not gaited
Colombian Criollo	Colombia	35	1	8	26	14.3	0.16	Not gaited
Trote y Galope								
Comtois	France	6	0	0	6	0.0	NA	Not gaited
Connemara Pony	Ireland	35	0	0	35	0.0	NA	Not gaited
Czech Warmblood	Czech Republic	19	0	0	19	0.0	NA	Not gaited
Dafur Pony	Sudan	14	0	0	14	0.0	NA	Not gaited
Dartmoor Pony	Great Britain	18	0	0	18	0.0	NA	Not gaited
Dølahest	Norway	16	0	0	16	0.0	NA	Not gaited
Don	Russia	16	0	0	16	0.0	NA	Not gaited
Dülmen Pony	Germany	7	0	0	7	0.0	NA	Not gaited
Exmoor	Great Britain	27	0	0	27	0.0	NA	Not gaited
Fell Pony	Great Britain	15	0	0	15	0.0	NA	Not gaited
Fjord	Norway	27	0	0	27	0.0	NA	Not gaited
French Saddle Horse	France	19	0	0	19	0.0	NA	Not gaited
Friesian	Netherlands	15	0	0	15	0.0	NA	Not gaited
Galiceno	Mexico and Spain	26	0	0	26	0.0	NA	Not gaited
German Riding Horse	Germany	11	0	0	11	0.0	NA	Not gaited
German Riding Pony	Germany	10	0	0	10	0.0	NA	Not gaited
Gotland Pony ⁴	Sweden	29	0	0	29	0.0	NA	Not gaited
Haflinger	Austria	18	0	0	18	0.0	NA	Not gaited
Hucul	Czech Republic and Poland	23	0	0	23	0.0	NA	Not gaited
Hungarian Coldblood	Hungary	11	0	0	11	0.0	NA	Not gaited
Kabardian	Russia (Caucasus)	10	0	0	10	0.0	NA	Not gaited
Khartoum Pony	Sudan	10	0	0	10	0.0	NA	Not gaited
Kinsky Horse	Czech Republic	19	0	0	19	0.0	NA	Not gaited
Kiso	Japan	24	0	0	24	0.0	NA	Not gaited
Kladruby ⁵	Czech Republic	37	0	0	37	0.0	NA	Not gaited
Knabstrup	Denmark	25	0	0	25	0.0	NA	Not gaited
Konik (Polish Primitive)	Poland	28	0	0	28	0.0	NA	Not gaited
Kurd	Iran	17	0	2	15	5.9	0.07	Not gaited
Kuznet Horse	Russia (Western Siberia)	11	0	0	11	0.0	NA	Not gaited
Latvian	Latvia	3	0	0	3	0.0	NA	Not gaited
Liebenthaler	Germany	20	0	0	20	0.0	NA	Not gaited
Lipizzaner	Slovenia and Austria	8	0	0	8	0.0	NA	Not gaited
Losino	Spain	10	0	0	10	0.0	NA	Not gaited
Lusitano	Portugal	19	0	0	19	0.0	NA	Not gaited
New Forest	Great Britain	42	1	6	35	9.5	1.23	Not gaited
Nordland	Norway	30	0	0	30	0.0	NA	Not gaited
Noriker	Austria	18	0	0	18	0.0	NA	Not gaited
North Swedish Horse ⁴	Sweden	34	0	0	34	0.0	NA	Not gaited
Paint	USA	49	0	8	41	8.2	0.39	Not gaited
Percheron	France	35	0	2	33	2.9	0.03	Not gaited
Polish Heavy Horse	Poland	14	0	0	14	0.0	NA	Not gaited
Potoka	Spain	10	0	0	10	0.0	NA	Not gaited
Przewalski's Horse ⁴	Asia	27	0	0	27	0.0	NA	Not gaited
Quarter Horse	USA	104	1	3	100	2.4	15.43***	Not gaited
Rhineland Heavy Draught	Germany	4	0	0	4	0.0	NA	Not gaited
Russian Riding Horse	Russia	14	0	0	14	0.0	NA	Not gaited
Shagya Arabian	Hungary	19	0	0	19	0.0	NA	Not gaited
Shetland Pony ⁴	Great Britain	55	0	0	55	0.0	NA	Not gaited
Shire	Great Britain	19	0	0	19	0.0	NA	Not gaited
Sorraia	Portugal	16	0	0	16	0.0	NA	Not gaited
South German Coldblood	Germany	3	0	0	3	0.0	NA	Not gaited
Suffolk	Great Britain	19	0	0	19	0.0	NA	Not gaited
Sumba	Indonesia	6	0	0	6	0.0	NA	Not gaited
Swedish Riding Pony	Sweden	2	0	0	2	0.0	NA	Not gaited
Swedish Warmblood ⁴	Sweden	64	0	0	64	0.0	NA	Not gaited

Table 1 (continued)

Breeds	Country of origin ¹	n	Genotype			Frequency of A allele (%)	Chi-square values for HWE test ²	Gaitedness ³
			AA	CA	CC			
Taishu	Japan	24	0	0	24	0.0	NA	Not gaited
Thessalian	Greece	10	1	3	6	25.0	NA	Not gaited
Thoroughbred ⁴	Great Britain	55	0	0	55	0.0	NA	Not gaited
Timor Pony	Indonesia (Timor)	6	0	0	6	0.0	NA	Not gaited
Tinker	Great Britain	11	0	0	11	0.0	NA	Not gaited
Trakehner	Germany	34	0	0	34	0.0	NA	Not gaited
Waler	Australia	15	0	0	15	0.0	NA	Not gaited
Welsh Mountain Pony	Great Britain	16	0	3	13	9.4	0.17	Not gaited
Welsh Pony	Great Britain	54	0	8	46	7.4	0.35	Not gaited
Yakut	Russia (Siberia)	25	0	0	25	0.0	NA	Not gaited
Yamut	Iran	17	0	1	16	2.9	0.02	Not gaited
Zemaitukai	Lithuania	15	0	0	15	0.0	NA	Not gaited
Coldblood Trotter	Sweden	306	50	176	80	45.1	7.98**	Harness – some pace
French Trotter ⁴	France	59	36	20	3	78.0	0.01	Harness
German Trotter	Germany	9	4	2	3	55.6	NA	Harness – some pace
Hackney Pony	England	35	1	1	33	4.3	14.87***	Harness
Orlov Trotter	Russia	5	0	2	3	20.0	NA	Harness – some pace
Standardbred Pacer ⁴	USA	40	40	0	0	100.0	NA	Harness – all pace
Standardbred Trotter ⁴	Sweden	270	253	17	0	96.9	0.29	Harness – some pace
Standardbred Trotter ⁴	USA	57	57	0	0	100.0	NA	Harness
Bashkir Pony	Russia	29	0	1	28	1.7	0.01	Unknown
Cheju	South Korea	14	0	4	10	14.3	0.39	Unknown
Israeli Local	Israel	9	0	1	8	5.6	NA	Unknown
Jordanian	Jordan	5	0	1	4	10.0	NA	Unknown
North Sudan Pony	Sudan	9	0	0	9	0.0	NA	Unknown
Retuertas	Spain	10	0	2	8	10.0	NA	Unknown
Tushuri Cxeni (Tushuri Horse)	Georgia	15	4	6	5	46.7	0.58	Unknown
Viatka	Russia	2	0	0	2	0.0	NA	Unknown

¹Country of origin refers to origin of the breed, not necessarily the origin of the samples.

²Chi-square values for tests of Hardy–Weinberg equilibrium (HWE) are given in breeds where applicable; degrees of freedom = 1.

³Summary of information on the degree and type of gaitedness that we have been able to collect based on breed descriptions. Gaited, the breed is generally considered gaited; some, a fraction of the horses of this breed are gaited, whereas others are not; not gaited, gaited horses are generally not observed for this breed; harness, breed used for harness racing; unknown, information on gait not found. *Gait descriptions*: lateral gaits, two beats: pace; lateral gaits, four beats (denoted in parenthesis are differences in speed, rhythm, liquidity of movement as specified by breed specific names): paso gaits (fino, corto, largo, llano, marcha), rack, revaal, running walk, slow gaits (stepping pace, singlefoot), tölt; diagonal gaits, four beats: foxtrot, trocha, marcha batida; ? means that gait classification is uncertain.

⁴All or some of the horses in this breed were included in the previous study by Andersson *et al.* (2012).

⁵The material included 23 Black Kladruby and 14 Grey Kladruby horses.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

22 967 656 bp) were genotyped; the latter SNP was the only SNP that showed a statistically significant association with the ability to pace in the previous genome-wide association analysis (Andersson *et al.* 2012). The *DMRT3_Ser301STOP* mutation was genotyped using a custom-designed TaqMan SNP Genotyping Assay (Applied Biosystems) as previously described (Andersson *et al.* 2012). The *BIEC2_620109* SNP was also genotyped with a TaqMan SNP Genotyping Assay that included the following reagents: forward primer, 5'-GCAAAGTGCAGAAATAGTCTTTTGGGA-3'; reverse primer, 5'-CACTCTTTTGGGAATGGTTCACATTAAGG-3'; wild-type probe (FAM), 5'-TAGTGCAAACGGTACGTT-3' and mutant probe (VIC), 5'-AAATAGTGCAAACAGTACGTT-3'. The amplification was carried out in a final volume of 5 μ l, with 2.5 μ l of Genotyping Master Mix (Applied Biosystems),

0.125 μ l Genotyping Assay, 0.875 μ l deionised water and 1.5 μ l DNA. Thermal cycling consisted of 95 °C for 10 min, followed by 50 cycles of 95 °C for 15 s and 60 °C for one min. For individuals genotyped in Uppsala, where TaqMan genotyping did not yield any results for the stop mutation ($n = 39$), and for some individuals, where it was deemed interesting to investigate whether they might possess any other *DMRT3* mutation, the *DMRT3* coding region was PCR-amplified and Sanger-sequenced as previously described (Andersson *et al.* 2012).

Data analysis

Sequences were edited using Codon Code Aligner (Codon-Code Corporation). Pairwise linkage disequilibrium between

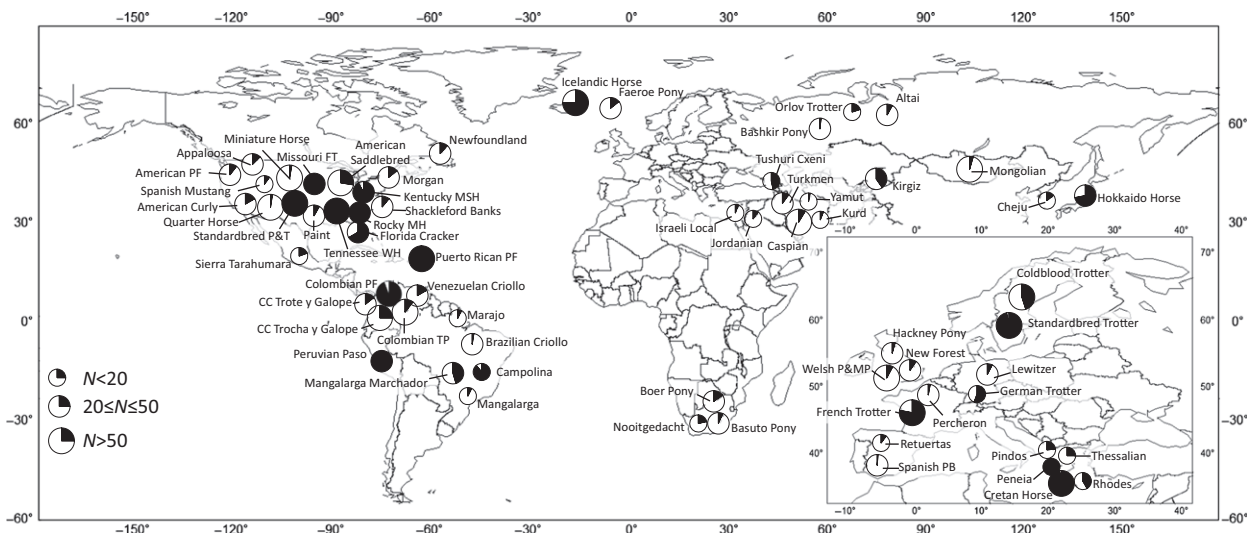


Figure 1 Frequency distribution of the *DMRT3* gait-altering mutation for breeds with at least five samples tested. The sizes of the pie-charts reflect sample size. Only breeds for which the mutation was observed are shown. The following abbreviations are used in the figure: CC, Colombian Criollo; FT, Fox Trotter; MH, Mountain Horse; MSH, Mountain Saddle Horse; PB, Pure Breed; PF, Paso Fino; P&MP, Pony and Mountain Pony; P&T, Pacer and Trotter; TP, Trocha Pura; WH, Walking Horse.

the *DMRT3_Ser301STOP* and *BIEC2_620109* was tested in ARLEQUIN 3.5.1.2 (Excoffier *et al.* 2005) using 1000 permutations. PHASE implemented in DNASP v5 (Librado & Rozas 2009) was used to reconstruct the gametic phase of the two SNPs, with default settings. Hardy–Weinberg equilibrium (HWE) for the *DMRT3_Ser301STOP* SNP was tested for in breeds that displayed both wild-type and mutant alleles and where sample size was at least 14 individuals (53 breeds). Allele frequencies were visualised as pie-chart graphs in Microsoft EXCEL with the background of the world map created in PANMAP (Diepenbroek *et al.* 2000).

Results

We generated *DMRT3_Ser301STOP* genotypes for a total of 4396 individuals belonging to 141 breeds (Table 1, Fig. 1). The number of samples per breed varied from one (British Riding Pony) to 306 (Swedish Coldblood Trotter), with an average of 31 samples per breed. The mutation was observed in horses from 68 breeds. Within these breeds, the frequency of the mutant allele (A) varied from 1.1% (Spanish Pure Breed, $n = 46$) to complete fixation in some North and South American gaited breeds (Table 1, Fig. 1). Nine of the tested breeds showed significant deviations from HWE (Table 1).

A subset of 2749 horses also was genotyped for SNP *BIEC2_620109*, which is located about 32 kb upstream of *DMRT3_Ser301STOP* (Andersson *et al.* 2012). The two SNPs are significantly linked according to the pairwise linkage disequilibrium test in ARLEQUIN ($\chi^2 = 5314$, $df = 2$, $P < 0.0001$). The genotype frequencies at these two SNPs were used to estimate the haplotype frequencies using PHASE

Table 2 Estimated haplotype frequencies for the *DMRT3_Ser301STOP* mutation (C>A) and the closely linked SNP *BIEC2_620109* (C>T) in a total of 2749 horses across different breeds.

Haplotypes	<i>BIEC2_620109</i>	<i>DMRT3_Ser301STOP</i>	n	Frequency
Wild type at both loci	C	C	3237	0.589
Mutant 620109 and wild-type <i>DMRT3</i>	T	C	109	0.020
Mutant 620109 and mutant <i>DMRT3</i>	T	A	2133	0.388
Wild-type 620109 and mutant <i>DMRT3</i>	C	A	19	0.003

n , number of chromosomes.

software (Table 2; Phase probabilities for all cases 1.00). This analysis revealed two predominant haplotypes across all breeds. C-C (i.e. the reference allele at both SNP loci) was the most common haplotype (59%; Table 2). T-A was the other abundant haplotype (39%), with the non-reference allele (T) at the *BIEC2_620109* SNP and the *DMRT3* nonsense mutation (A). The third most common haplotype (2%) was T-C, with the non-reference allele (T) at the *BIEC2_620109* SNP and the reference allele (C) at *DMRT3*. The simplest interpretation of these data is that T-C represents the ancestral haplotype on which the *DMRT3* nonsense mutation arose and that the T-A haplotype has increased in frequency because of strong positive selection for the gait mutation. This interpretation is supported by the observation that the T-C haplotype was observed in 16 breeds that lack the *DMRT3* nonsense mutation (Table S1), whereas no breed displayed only the *DMRT3* nonsense mutation in the absence of the *BIEC2_620109* SNP

non-reference (T) allele. The fourth haplotype (C-A) was found in the heterozygous state in only 19 horses, representing six breeds, and of 5498 haplotypes in total (0.3%, Table 2, Table S1). This haplotype has most likely originated through recombination between the C-C and T-A haplotypes.

Discussion

The present study has demonstrated that the *DMRT3* gait-altering nonsense mutation has a worldwide distribution. It was detected in 68 of 141 breeds included in the study. The very strong linkage disequilibrium between this polymorphism and the closely linked *BIEC2_620109* SNP across all tested breeds implies that this mutation arose once and has spread across the world by positive selection. The positive selection is most certainly related to the appreciation of the smooth ride offered by ambling gaits as well as the increased performance capacity of mutant horses when raced in gaits other than gallop. The mutation is abundant in all breeds classified as gaited, for instance the Hokkaido Horse and the Icelandic Horse, and in horses used for harness racing. However, the mutation also occurs at a low frequency in many other breeds not classified as gaited, such as the Welsh and New Forest Ponies from the UK. The presence of the *DMRT3* mutation in breeds that are not generally classified as gaited might be explained by a preference for ambling gaits in the past, by historical crossing with gaited breeds or by the mutation having segregated in the ancestral population prior to breed formation. Cross-breeding has been used frequently to improve vigour of breeds that were nearly extinct and suffered from inbreeding, or to introduce a desired conformation, character or colour trait. One such example might be the Israeli Local horse, which has been influenced by the Tennessee Walking Horse amongst other breeds (Hendricks 1995) and in which we observed the mutation at a frequency of 5.6%. The Miniature Horse is a breed with an influence of many different breeds in the past, and in our data set, the *DMRT3* mutation occurs at a frequency of 3.1% (Table 1). Similarly, the Bashkir Pony (allele frequency 1.7%) is known to have been cross-bred to, amongst others, trotters (Hendricks 1995). If gaitedness was not disadvantageous in such populations, one can surmise that there was little or no selection applied to eliminate it, and thus, the mutation could persist in the population at low frequency. Similarly, little or no selection would be applied if the phenotypic effect of the mutation is not obvious, as may be the case in the heterozygous state. It will be of considerable interest to study genotype–phenotype relationships between *DMRT3* mutant and wild-type horses from such ‘non-gaited’ breeds.

We observed a significant deviation from expected Hardy–Weinberg proportions in nine of the breeds included in this study, which is more than expected by chance. There

are many possible reasons why significant deviations from HWE may occur, including non-random mating, population substructure, selection and cross-breeding, but may also be due to genotyping errors. We are convinced that the deviation from HWE in this study is not due to genotyping errors but rather reflects selection and non-random mating given that the *Gait keeper* mutation has a major effect on a crucial trait in the domestic horse.

All gaited horse breeds that we tested carry the *DMRT3_Ser301STOP* mutation, and many of them at a very high frequency or in complete fixation. However, there are exceptions where individual horses claimed to be gaited have tested negative. The best documented cases concern Icelandic horses for which at least some horses homozygous C-C for the wild-type allele can amble (tölt) to a variable extent, but are unable to pace. Another mutation in the coding sequence of *DMRT3* has been excluded for these individuals, which leaves the possibility of (i) a high-impact *DMRT3* mutation outside the coding sequence, (ii) a high-impact mutation affecting another gene important to regulation of locomotion or (iii) a polygenic effect. There has been strong selection over many generations in the Icelandic Horse for the ability to amble with perfect smoothness. It is possible that this strong selection pressure led to enrichment for mutations promoting ambling capacity so that even horses that are homozygous wild type at *DMRT3* can amble, with good training and riding. Thus, an important topic for future research is to screen for additional loci affecting gaitedness in horses. In fact, the discovery of the *DMRT3* mutation in horses shows that the domestic horse is probably the best animal model available to study the genetic control of locomotion because of the very long history of selection for different types of locomotion (speed in gallop, trot and pace, high jumping and dressage).

It has long been discussed whether gaitedness is a trait that the domestic horse inherited from its ancestors or acquired during the process of domestication. In support of the former hypothesis, a study on footfall patterns of three individuals of *Hipparion* (dated to 3.5 Myr ago) found that they were moving at a running walk, equivalent to the tölt of the Icelandic Horses (Renders 1984). However, gaitedness in domestic horses must be a derived trait given that it is so strongly associated with the *DMRT3_Ser301STOP* mutation. It is clear that this is a recent, derived mutation because all tested Przewalski’s Horses, a close relative to the wild ancestor of the domestic horse, were homozygous wild type at *DMRT3*. We also extracted sequence data from two recently published genome sequences of ancient horses (Orlando *et al.* 2013), and only the wild-type sequence was found. Furthermore, the mutation disrupts the *DMRT3* coding sequence that is otherwise maintained in all vertebrates for which sequence information is available (Andersson *et al.* 2012).

Our current characterisation of the haplotype associated with the *DMRT3_Ser301STOP* mutation does not have the

resolution required to firmly conclude when and where the mutation arose. This will require re-sequencing of the 440 kb region showing strong linkage disequilibrium with the gait mutation (Andersson *et al.* 2012) to identify diagnostic polymorphisms that were present on the ancestral haplotype or that have occurred on the mutant haplotype subsequent to the causal mutations. The present study provides an excellent screening for horses to be used in such a detailed haplotype analysis. Tracing how humans have spread the *DMRT3* mutation across the globe using modern samples is challenging because the frequency of the *DMRT3* mutation is strongly influenced by the preferred use of horses. For instance, Icelandic horses originate primarily from Scandinavian horses brought by the settlers in the 9th century, and the Icelandic sagas strongly indicate that horses with the ability to pace were present on Iceland 1100 years ago (Íslendingasagnautgáfan 1946). However, we did not find the *DMRT3* mutation in contemporary horses representing breeds originating from Scandinavia (e.g. Fjord, Gotland Pony and North Swedish Horse), with the exception of Coldblood Trotters used for harness racing. A possible explanation is that the way horses are used in Scandinavia has changed since the 9th century when Iceland was settled and that Scandinavian horses were for a long period of time more extensively used as draught horses in agriculture and forestry, which was not the case on Iceland. A general trend in our analysis across breeds is that the *DMRT3* mutation is rare or absent in draught horse breeds. Furthermore, the *DMRT3* mutation is very common in horse breeds from both South and North America, particularly in the USA, as proved by several of the American breeds showing complete fixation of the mutant allele (Fig. 1, Table 1). All American breeds are derived from horses originating from Europe, particularly the Iberian Peninsula (Rodero *et al.* 1992; Vega-Pla *et al.* 2005; Luis *et al.* 2006). Although we observed the mutation in Spanish horses, it was present at a very low frequency. Nevertheless, we did not have any samples of the extinct jennet type, which might have been the type of horse first taken to the Western Hemisphere by the early colonisers and which was known to be gaited (Hendricks 1995). Thus, a possible reason for the low frequency of the gait mutation in modern Spanish horses is that ambling is considered a negative trait in most Spanish breeds because the use of horses has changed since the time Spanish conquistadores arrived in the Americas.

To conclude, the extensive worldwide sampling shows that the *DMRT3* Gait keeper mutation is not restricted to a geographical range but has spread across the world. In general, it follows the distribution of breeds classified as gaited and breeds used for harness racing, but it also appears at various frequencies in other breeds. It is still unclear where the mutation arose. A large-scale analysis of the whole haplotype should shed more light on this intriguing question.

Acknowledgements

We most sincerely thank all sample contributors: Jennifer Leonard (Retuertas), Yekrun Horse Industry Development Co. Ltd. (horses from Iran), Dr. T. Ohnuma (Hokkaido Horses), Lutfi Musa (Sudanese breeds), Sonia Zakizadeh (Yamut and Turkmen), Sergey Knyasev (Altai and Kuznet Horse), Ines Wemken (Barb horses), Julia McCann (American Saddlebred Horses) and Mario Calcagno (Marwari). We are grateful to Dr. Deb Bennett for help with gait classification, Reto Burri for discussions and Markus Sällman-Almén for extracting published sequence data from ancient horses. The study was financially supported by the Knut and Alice Wallenberg Foundation.

Conflict of interest

Leif Andersson, Lisa Andersson and Gabriella Lindgren are co-inventors on a patent application concerning commercial testing of the *DMRT3* mutation.

References

- Andersson L.S., Larhammar M., Memic F. *et al.* (2012) Mutations in *DMRT3* affect locomotion in horses and spinal circuit function in mice. *Nature* **488**, 642–6.
- Bateson W. (1907) Trotting and pacing: dominant and recessive? *Science* **26**, 908.
- Diepenbroek M., Grobe H. & Sieger R. (2000) PANMAP, <http://www.pangaea.de/Software/PanMap>.
- Excoffier L., Laval G. & Schneider S. (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.
- Hendricks B. (1995) *International Encyclopedia of Horse Breeds*. University of Oklahoma Press, Norman, OK, USA.
- Hong C.S., Park B.Y. & Saint-Jeannet J.P. (2007) The function of *Dmrt* genes in vertebrate development: it is not just about sex. *Developmental Biology* **310**, 1–9.
- Íslendingasagnautgáfan (1946) *Íslendinga sögur*. Hrapseyjarprent hf. Reykjavík, Iceland.
- Librado P. & Rozas J. (2009) DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–2.
- Luis C., Bastos-Silveira C., Cothran E.G. & Oom M.D. (2006) Iberian origin of New World horse breeds. *Journal of Heredity* **97**, 107–13.
- Nicodemus M.C. & Clayton H.M. (2003) Temporal variables of four-beat, stepping gaits of gaited horses. *Applied Animal Behaviour Science* **80**, 133–42.
- Orlando L., Ginolhac A., Zhang G. *et al.* (2013) Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–8.
- Raymond C.S., Shamu C.E., Shen M.M., Seifert K.J., Hirsch B., Hodgkin J. & Zarkower D. (1998) Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691–5.
- Renders E. (1984) The gait of *Hipparion* sp. from fossil footprints in Laetoli, Tanzania. *Nature* **308**, 179–81.
- Robilliard J.J., Pfau T. & Wilson A.M. (2007) Gait characterisation and classification in horses. *Journal of Experimental Biology* **210**, 187–97.

- Rodero A., Delgado J.V. & Rodero E. (1992) Primitive Andalusian livestock and their implications in the discovery of America. *Archivos de Zootecnia* **41**, 383–400.
- Vega-Pla J.L., Calderón J., Rodríguez-Gallardo P.P., Alcaide B., Sereno F.T.P.S., Costa M.R., Pérez-Pineda E., Martínez A.M., Delgado J.V. & Rico C. (2005). The Retuertas horse: the “missing link” in the Iberoamerican horse breeds origin? In: *Conservation Genetics of Endangered Horse Breeds*. (Ed. by I. Bodó, L. Alderson & B. Langlois), pp. 167–76. EAAP No. 116. Wageningen Academic Publishers, Wageningen, The Netherlands.

Supporting information

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

Table S1. Occurrence of the two rare haplotypes mutant (T) – wild-type (C) and wild-type (C) – mutant (A) for SNP *BIEC2_620109* and *DMRT3_Ser301STOP* respectively across horse breeds.