A single-nucleotide mutation within the **TBX3** enhancer increased body size in Chinese horses

**Highlights**

- One single A/G SNP in TBX3 enhancer region drives size variation in Chinese horses
- The frequency of the G variant correlates positively with size in 763 horses
- Cellular and mice models confirm it affects TBX3 transcription and the limb length
- The G variant first occurred ~2,300 years ago and rose in frequency since

**Authors**

Xuexue Liu, Yanli Zhang, Wujun Liu, ..., Ludovic Orlando, Yuehui Ma, Lin Jiang

**Correspondence**

ludovic.orlando@univ-tlse3.fr (L.O.), yuehui.ma@263.net (Y.M.), jianglin@caas.cn (L.J.)

**In brief**

Liu et al. discover a regulatory mutation in TBX3 driving size variation in Chinese horses. Reporter assays and mice models confirm that it can increase TBX3 transcription and the limb length. They find this mutation first occurred ~2,300 years ago and rose in frequency since, revealing size as one of the main selection targets of past Chinese breeders.
A single-nucleotide mutation within the TBX3 enhancer increased body size in Chinese horses

Xuexue Liu,1,2,3,6 Yanli Zhang,1,3,6 Wujun Liu,4 Yefang Li,1,3 Jianfei Pan,1,3 Yabin Pu,1,3 Jianlin Han,3,5 Ludovic Orlando,2,* Yuehui Ma,1,3,* and Lin Jiang1,3,8,*

1Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, P.R. China
2Centre d’Anthropobiologie et de Génomique de Toulouse, Université Paul Sabatier, 37 allées Jules Guesde, 31000 Toulouse, France
3CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, P.R. China
4College of Animal Science, Xinjiang Agriculture University, Urumqi, Xinjiang, China
5International Livestock Research Institute (ILRI), Nairobi 00100, Kenya
6These authors contributed equally
7Twitter: @Xuexue97313824
8Lead contact
*Correspondence: ludovic.orlando@univ-tlse3.fr (L.O.), yuehui.ma@263.net (Y.M.), jianglin@caas.cn (L.J.)

https://doi.org/10.1016/j.cub.2021.11.052

SUMMARY

Chinese ponies are endemic to the mountainous areas of southwestern China and were first reported in the archaeological record at the Royal Tomb of Zhongshan King, Mancheng, dated to approximately ~2,100 YBP.1 Previous work has started uncovering the genetic basis of size variation in western ponies and horses, revealing a limited number of loci, including HMGA2,2 LCORL/NCAPG,3 ZFAT, and LASP1.4,5 Whether the same genetic pathways also drive the small body size of Chinese ponies, which show striking anatomical differences to Shetland ponies,6 remains unclear.2,7 To test this, we combined whole-genome sequences of 187 horses across China. Statistical analyses revealed top association between genetic variation at the T-box transcription factor 3 (TBX3) and the body size. Fine-scale analysis across an extended population of 189 ponies and 574 horses narrowed down the association to one A/G SNP at an enhancer region upstream of the TBX3 (ECA8:20,644,555, p = 2.34e−39). Luciferase assays confirmed the single-nucleotide G mutation upregulating TBX3 expression, and enhancer-knockout mice exhibited shorter limbs than wild-type littermates (p < 0.01). Re-analysis of ancient DNA data showed that the G allele, which is most frequent in modern horses, first occurred some ~2,300 years ago and rose in frequency since. This supports selection for larger size in Asia from approximately the beginning of the Chinese Empire. Overall, this study characterized the causal regulatory mutation underlying small body size in Chinese ponies and revealed size as one of the main selection targets of past Chinese breeders.

RESULTS AND DISCUSSION

Genome sequencing and phylogenetic analysis

We sequenced the genomes of 48 ponies and nine horses at an average depth of coverage of ~11.1x and combined them with 130 genomes previously characterized7,8 (Data S1A; STAR Methods). Our dataset includes 52 Chinese ponies from four breeds, 100 Chinese native horses from 12 breeds, 23 imported Falabella ponies, seven Thoroughbred horses, and five Przewalski’s horses. Stringent quality filtering identified a total of 20,428,318 SNPs. Two main clusters consisting of Northern horses (NC) and Qinghai-Tibetan horses (QT) were identified, in line with previous analyses.8 Pony breeds showed genome-wide heterozygosity levels generally higher than those found in horse breeds (Data S1C; STAR Methods). Their genomes were also characterized by shorter runs of homozygosity fragments (ROHs) (Figure S1D), possibly reflecting relaxed selective pressures in Chinese ponies relative to horses.

Selection signatures among horses and ponies

We scanned the genomes of 152 Chinese native ponies and horses for signatures of positive selection that may underpin their different statures. To achieve this, we calculated three complementary statistics along the horse reference genome using 100-kb-long sliding windows (step size = 15 kb). The first statistic was the pairwise genetic differentiation FST index for identifying those genomic regions maximizing allelic frequency differences between ponies and horses. The second statistic measured relative differences in nucleotide diversity between the two groups (θπ(pony/horse)), with high or (low) θπ values indicative of positive selection in Chinese horses or ponies, respectively. Selection
candidates were also identified through the calculation of cross-population extended haplotype homozygosity (XP-EHH). In order to limit false positive identifications, we considered the top 1% windows from all three scans, which provided a total of 153 windows encompassing 4.66 Mb and 69 protein-coding genes (Figure 1A; Data S2A–S2D; STAR Methods). These selection candidates showed limited genetic diversity in horses, but not in ponies, suggesting positive selection for horse-related phenotypes (Figure 1A). Most of the protein-coding genes were significantly enriched for functional Gene Ontology categories related to the anterior/posterior pattern specification (p = 7.77e-7, following g:SCS “Set Counts and Sizes” correction), the embryonic skeletal system development (p = 3.07e-6), embryonic skeletal system morphogenesis (p = 1.27e-5), and pattern specification process (adjusted p = 3.59e-5) (Data S2E). Importantly, all the four outlier genomic regions providing the strongest signatures of selection contained TBX3 (ECA8:20,640,001–20,740,000 bp, ECA8:20,655,001–20,755,000 bp, ECA8:20,625,001–20,725,000 bp, and ECA8:20,610,001–20,710,000 bp) (Figure 1A; Data S2D) and were further refined into an 80-kb region (ECA8:20,640,000–20,720,000). Red dots indicate SNPs reaching statistical significance are shown in red. Two SNPs (ECA8:20,644,555 and ECA8:20,644,525) overlapping the GERP and enhancer regions are labeled in the red box, together with the annotated regions.

See also Figure S1 and Data S2A–S2P.

TBX3 is a T-box gene involved in the development of the anterior/posterior axis, which was previously reported to undergo selection in Debao ponies. This indicated that TBX3 could play an important role in driving body size differences between Chinese ponies and horses. Interestingly, the strongest selection signature identified in Falabella ponies, which represent the smallest miniature horse breed on the planet, was located around another gene, HMGA2 (Figure S2A), which was previously described to cause asymmetric deposition of skin pigments characteristic of Dun phenotypes when present in the ancestral, non-deleted allelic state. In order to check whether the selection signature identified above could in fact target Dun phenotypes, we screened the presence of not deleted (Dun and non-dun1) and deleted (non-dun2) alleles in all 187 genomes present in our panel (Figure S2B; Data S2L). We found Dun alleles in five Przewalski’s horses (PrZ) and one Chinese Hequ horse, but not in any of the remaining 99 Chinese horses, 52 Chinese ponies, 23 Falabella ponies, and 7 Thoroughbred horses. This ruled out Dun phenotypes driving the selection signatures identified.
In order to further identify the main underlying genetic driver, we investigated patterns of differential placentome gene expression in four Chinese YiLi horses and three NiQi ponies (Figure S2C). We focused on the placenta as this tissue is essential for fetal growth, and the placentome surface structure is known to be associated with the neonatal body size. Overall, we identified a total of 620 differentially expressed genes in the transcriptomes of YiLi horses and NiQi ponies (Data S2N). Interestingly, TBX3 was highly expressed in horse placentas. In contrast, none of the other two candidate genes (TBX5 and HMGA2) showed substantial expression levels (fragments per kilobase of exon per million fragments mapped, FPKM < 0.3). TBX3 placentome expression was also 4.3-fold higher in YiLi horses than NiQi ponies (p = 0.04; Figure 1E), indicating positive association of TBX3 expression levels and adult wither height. This finding was in line with the absence of missense mutations in the TBX3 region examined, and the lower nucleotide diversity and negative Tajima’s D statistics found in this region in horses (Figure 1B). Altogether, this suggested that the genetic variation present within the TBX3 region could regulate gene expression and potentially drive size differences in Chinese horses and ponies. We thus decided to focus the following experimental validation efforts on the TBX3 locus.

**Identifying TBX3 causative variant**

As the TBX3 region showing selection signatures and associated with wither height encompassed approximately ~80 kb, we carried out a number of analyses to narrow down the list of possible causative variants (STAR Methods). We first explored the patterns of linkage disequilibrium (LD) in the region using a haplotype block analysis (Data S2O), which revealed nine blocks (H1–H9, haplotype p < 1e-24) containing 16 candidate SNPs (Figure 1F). Of these, only the H2 block, which overlapped two SNPs (TBX3-EN1, ECA8:20,644,525 and TBX3-EN2, ECA8:20,644,555), overlapped with a 243-bp region showing extremely elevated conservation GERP scores in 90 mammals (score = 504.5, p = 1.51e-166). These two SNPs also overlapped an enhancer region previously identified in large-scale mouse embryonic studies (Figure 1F). Together, this suggested that the TBX3-EN1 and TBX3-EN2 SNPs may play a strong role in the regulation of TBX3 expression (Figure 1F; Data S2P).

We next measured the allelic frequency at the TBX3-EN1 and TBX3-EN2 SNPs in a comprehensive panel of 189 Chinese ponies and 574 Chinese horses using competitive allele specific PCR (KASP) assays (Data S3A–S3C; STAR Methods). Chi-square tests confirmed strong statistical association with the horse wither height, which was stronger for TBX3-EN2 (p = 2.34e-39) than TBX3-EN1 (p = 8.18e-29) (Figure 2). Our association model suggested that the variation at TBX3-EN2 (TBX3-EN1) could explain as much as ~20.3% (~15.1%) of the height variation (~0.1 cm) in this panel (Figure 2A).

Interestingly, the A allele at TBX3-EN2 was most common in ponies belonging to seven pony breeds (N = 189; JiCh, DeBa, BaSe, GuZh, YuNa, LCh, and JiJi), one Przewalski’s horse breed (N = 13), and 18 horse breeds (N = 561; WuS, JiZi, LKZ, NiMu, MZH, BLK, MoGo, WuZh, DaTo, HSK, YaQi, CDM, SeSe, HeQu, XNH, SaHe, GZg, and YiLi). (A) Allelic frequencies at the TBX3-EN1 (ECA8:20,644,525) locus. The A allele is in blue and the G allele is in orange. (B) Allelic frequencies at the TBX3-EN2 (ECA8:20,644,555) locus (left y axis). The A allele is in blue and the T allele is in orange. TBX3-EN2 (p = 2.34e-39) and TBX3-EN1 (p = 8.18e-29) showed strong correlation with horse wither height, explaining 20.27% and 15.05% of the phenotypic variation, respectively. The yellow and green bars represented the pony and horse. Average wither’s height within breed is indicated by the black line. Cartoon pictures of ponies and horses were generated according to the photos from the Animal Genetic Resources in China: horse, donkey, and camel (2011). See also Data S3A and S3B.

**Genome-wide association study**

We next carried out a genome-wide association analysis for the wither height of all 152 Chinese samples (STAR Methods). This analysis accounted for the underlying population structure by adjustment of above mentioned PCA covariates and genetic relatedness among all pairs of individuals (STAR Methods). It revealed two statistically significant regions (p < 6e−8), encompassing 82 SNPs (Figures 1C and 1D; Data S2M). The first region spanned the TBX2/TBX5 locus (ECA8:20,566,781–20,865,109), which also provided the top-selection candidate. The second region comprised the HMGA2 locus ECA6:82,576,739–82,701,867. This result indicated that TBX3, TBX5, and/or HMGA2 may provide the genetic basis for the short stature of Chinese ponies.
breeds (71.4%–90.0%, wither height > 135 cm). Intermediate allelic frequencies were found in Tibetan horses (67.2%–87.5%), which shows an intermediate wither height (125–135 cm) (Figure 2A; Data S3B). This confirmed a model in which the A variant could represent the ancestral allele at TBX3-EN2 and the derived G allele increased in frequency among Chinese horses in response to selection for higher wither height. The distribution of the genetic variation at TBX3-EN1 in our panel suggested a possibly similar scenario for the T allele.

We further reconstructed the frequency trajectories for the TBX3-EN1 and TBX3-EN2 derived alleles during the last ~6,000 years, leveraging an extensive ancient genome time series previously released for horses.19 We found that most ancient samples carried the A allele at TBX3-EN2 until ~2,300 YBP, a time after which the G allele increased in frequency steadily until the present time (Figures 3A and 3B; STAR Methods). Importantly, post-mortem DNA decay can lead to the accumulation of G>A substitutions in ancient DNA data.20 The allelic trajectory reconstructed here was, however, not correlated to post-mortem DNA damage levels, as measured through mapDamage2 deamination parameters21 (Figures 3B and S3). Therefore, our analyses supported ongoing selection for the G allele at TBX3-EN2 in relation with increased wither height from the late first millennium BCE. The T allele present at the TBX3-EN1 also strongly associated with horse wither height and was randomly distributed at the beginning but followed a similar trajectory with the increase of the TBX3-EN2 G frequency, in line with the selection acting on linked genetic variants (Figure 3B).

Moreover, the genotyping profiles of 245 ancient horses19 within the 10-kb region flanking the TBX3 enhancer containing TBX3-EN1 and TBX3-EN2 were closer to those found in Chinese ponies (N = 52) than in Chinese horses (N = 100) (Figure 3C). Modern Chinese horses were characterized by limited sequence variation, in line with the positive selection acting in the region.

Functional validation of the TBX3 causative variant

The analyses based on ancient and modern genomic data supported a model in which the selection for the T mutation at TBX3-EN1 and G mutation at TBX3-EN2 may have driven increased sizes in Chinese horses over the last 2,300 years. Reporter assays in HEK293T cells, however, indicated no significant difference in expression levels between the vector carrying both derived alleles at TBX3-EN1 and TBX3-EN2 (T and G), and that carrying the derived allele at TBX3-EN2 only (C and G; Figure 4A). This contrasted with the significant drop detected for vectors carrying the ancestral A allele at TBX3-EN2, regardless of the allele associated at TBX3-EN1 (T and A or C and A). Together, this indicated that the G mutation at TBX3-EN2 alone was sufficient to increase TBX3 expression levels and, thus, likely represented the causative allele. We then speculated that the G allele at TBX3-EN2 could affect the TBX3 enhancer region and TBX3 expression and may have driven larger wither height in Chinese horses during the last ~2,300 YBP. Interestingly, the horses examined for placental transcriptome expression supported such a model, as all four Yili horses were homozygous G at TBX3-EN2 while all three NiQiang ponies were homozygous A at TBX3-EN2 (Figure 1E).
In order to further validate the functional consequences of \( \text{TBX3} \) mutations, we constructed knockout (KO) mice for the \( \text{TBX3} \) enhancer using the CRISPR/Cas9 technology (STAR Methods), in which a core region showing high sequence similarity among vertebrates and including the \( \text{TBX3-EN2 SNP} \) was deleted (Data S4A; Figure 4B). Such \( \text{TBX3} \) enhancer \( ^{\text{−/−}} \) mice showed a 6-fold reduction in \( \text{TBX3} \) forelimb transcription relative to the WT (\( p = 0.004 \)). Similar reduction in \( \text{TBX3} \) transcription levels was also detected in embryos and 2-week leg bones (\( p < 0.05 \)) (Figures 4C and S4A). No differences in \( \text{TBX3} \) transcription were detected in muscle, skin, or heart tissues. Our results thus supported that the 35-bp region acts as a limb-specific enhancer for \( \text{TBX3} \) expression, in line with dual-luciferase reporter experiments in the HEK293T cell line (Figure 4A).

Finally, we used micro-CT and X-ray imaging techniques on 11- and 2-week mice to measure the length of all bone structures involved in limb formation (i.e., the humerus, radius, metacarpal, and fingers for forelimbs and the femur, tibia, metatarsal, and toes for hindlimbs) (Figure S4B). The 11-week mice included the littermate of ten males (WT:KO = 3:3) and six females (WT:KO = 5:3), while 2-week mice include the littermate of ten males (WT:KO = 3:7) and eight females (WT:KO = 5:3). KO and WT mice showed statistically significant differences in the length of all measured limb bones, with size reduction in KO mice ranging from 1.06% to 21.42% (\( p < 0.05 \)). This finding was consistent across sexes and the two developmental stages investigated (Data S4B–S4D; Figures 4D, 4E, and S4C). Interestingly, the size reduction was more pronounced in forelimbs than hindlimbs, especially in 2-week female mice, reaching almost 20% for the third and fourth metacarpal digits (Figures 4E and S4C). This was in agreement with previous studies reporting a crucial role in driving body size variation in Chinese horses and ponies. Interestingly, the A allele appeared ubiquitous among a crucial role in driving body size variation in Chinese horses and ponies. Interestingly, the A allele appeared ubiquitous among
ancient horses predating the Iron Age in the most extensive ancient genome time series available for horses. The G allele was first detected among Berel’ Pazyryk horses some ~2,300 YBP in the Altay region and steadily increased to present-day frequencies thereafter. This supported a scenario in which past breeders started to select horses of larger sizes concurrently with the beginning of the Chinese Empire during the Qin dynasty. Importantly, all Przewalski’s horses investigated carried the A ancestral allele, which was also predominant among Chinese ponies. Przewalski’s horses and Chinese pony breeds were, thus, not subject to selection for size during their recent evolutionary history. As the G allele, most frequent in horses, is also dominant in western ponies and miniature horses, including Falabellas, our work unveiled different mechanisms underlying size shifts in the history of horse breeding. It also offered a new genetic marker to the horse breeding industry for size improvement and/or pony conservation.

AUTHOR CONTRIBUTIONS
Conceptualization, L.J. and Y.M.; methodology, L.J., Y.M., and L.O.; investigation, X.L., Y.Z., J.P., Y.L., Y.P., and W.L.; writing – original draft, X.L., L.J., and L.O.; writing – review & editing, L.O., L.J., X.L., and J.H.; funding acquisition, L.J., Y.M., X.L., and L.O.; resources, Y.M., L.J., W.L., and Y.P.; supervision, L.J., Y.M., and J.H.

DECLARATION OF INTERESTS
The authors have a patent related to this work.

REFERENCES
1. Hou, W. (1990). A brief analysis of the origin and development of the pony in China. J. Agric. Archaeol. 1, 340–344.
2. Kader, A., Li, Y., Dong, K., Irwin, D.M., Zhao, Q., He, X., Liu, J., Pu, Y., Gorkhali, N.A., Liu, X., et al. (2019). Population variation reveals independent selection toward small body size in Chinese Debao pony. Genome Biol. Evol. 8, 42–50.
3. Metzger, J., Schrimpf, R., Philipp, U., and Distl, O. (2013). Expression levels of LCORL are associated with body size in horses. PLoS ONE 8, e56497.
4. Petersen, J.L., Mickelson, J.R., Rendahl, A.K., Valberg, S.J., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., et al. (2013). Genome-wide analysis reveals selection for important traits in domestic horse breeds. PLoS Genet. 9, e1003211.
5. Orr, N., Back, W., Gu, J., Leegwater, P., Govindarajan, P., Conroy, J., Ducro, B., Van Arendonk, J.A., MacHugh, D.E., Ennis, S., et al. (2010). Genome-wide SNP association-based localization of a dwarfism gene in Friesian dwarf horses. Anim. Genet. 41 (Suppl 2), 2–7.
6. Metzger, J., Gast, A.C., Schrimpf, R., Rau, J., Eikelberg, D., Beineke, A., Heilige, M., and Distl, O. (2017). Whole-genome sequencing reveals a potential causal mutation for dwarfism in the Miniature Shetland pony. Mamm. Genome 28, 143–151.
7. Asadollahpour Nanaei, H., Esmaillizadeh, A., Ayatollahi Mehrangi, A., Han, J., Wu, D.D., Li, Y., and Zhang, Y.P. (2020). Comparative population genomic analysis uncovers novel genomic footprints and genes associated with small body size in Chinese pony. BMC Genomics 21, 496.
8. Liu, X., Zhang, Y., Li, Y., Fan, J., Wang, D., Chen, W., Zheng, Z., He, X., Zhao, Q., Pu, Y., et al. (2019). EPAS1 gain-of-function mutation contributes to high-altitude adaptation in Tibetan horses. Mol. Biol. Evol. 2, 2591–2603.
9. Rautvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). gProfile: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 47 (W1), W191–W198.
10. Bamshad, M., Le, T., Watkins, W.S., Dixon, M.E., Kramer, B.E., Roeder, A.D., Carey, J.C., Root, S., Schinzel, A., Van Maldergem, L., et al. (1999). The spectrum of mutations in TBX3: genotype/phenotype relationship in uinar-mammary syndrome. Am. J. Hum. Genet. 64, 1550–1562.
11. Hendricks, B.L. (2007). International Encyclopedia of Horse Breeds (University of Oklahoma Press).
12. Frischknecht, M., Jagannathan, V., Platted, P., Neuditschko, M., Signer, Hasler, H., Bachmann, I., Pacholewska, A., Drögemüller, C., Dietschi, E., Flury, C., et al. (2015). A non-synonymous HMGAA2 variant decreases height in Shetland ponies and other small horses. PLoS ONE 10, e0140749.
13. Salek Ardestani, S., Aminafshar, M., Zandi Baghche Maryam, M.B., Banabazi, M.H., Sargolzaei, M., and Mier, Y. (2019). Whole-genome
signatures of selection in sport horses revealed selection footprints related to musculoskeletal system development processes. Animals (Basel) 10, 53.

14. Klecel, W., and Martyniuk, E. (2021). From the Eurasian Steppes to the Roman circuses: a review of early development of horse breeding and management. Animals (Basel) 11, 1859.

15. Imsland, F., McGowan, K., Rubin, C.-J., Henegar, C., Sundström, E., Berglund, J., Schwochow, D., Gustafson, U., Imsland, P., Lindblad-Toh, K., et al. (2016). Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. Nat. Genet. 48, 152–158.

16. Jansson, T., and Powell, T.L. (2007). Role of the placenta in fetal programming: underlying mechanisms and potential interventional approaches. Clin. Sci. (Lond.) 113, 1–13.

17. Alwasel, S.H., Abotalib, Z., Aljarallah, J.S., Osmond, C., Al Omar, S.Y., Harrath, A., Thornburg, K., and Barker, D.J.P. (2012). The breadth of the placental surface but not the length is associated with body size at birth. Placenta 33, 619–622.

18. Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. Nucleic Acids Res. 35, DB88–D92.

19. Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Paabo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA 104, 14179–14184.

20. Alwasel, S.H., Aljarallah, J.S., AlOmair, S.Y., Harrath, A., and Thornburg, K., et al. (2019). Tracking five millennia of horse management with extensive ancient genome time series. Cell 177, 1419–1435.e31.

21. Alwasel, S.H., Abotalib, Z., Aljarallah, J.S., Osmond, C., Al Omar, S.Y., Harrath, A., Thornburg, K., and Barker, D.J.P. (2012). The breadth of the placental surface but not the length is associated with body size at birth. Placenta 33, 619–622.

22. Bamshad, M., Lin, R.C., Law, D.J., Watkins, W.C., Krakowiak, P.A., Berglund, J., Schwochow, D., Gustafson, U., Imsland, P., Lindblad-Toh, K., et al. (2016). Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. Nat. Genet. 48, 152–158.

23. Zimmerli, D., Borrelli, C., Jauregi-Miguel, A., Söderholm, S., Brütsch, S., Douglass, N.F., et al. (2018). Improved reference genome for the domestic horse increases assembly contiguity and composition. Commun Biol 1, 104.

24. Yue, L., Honghai, C., Jing, Y., and Zhen, W. (2014). The horse bones unearthed from Zaoshugounao site in Chunhua County of Shaanxi Province. Journal of Northwest University 44, 311–317.

25. Yue, L., Honghai, C., Jing, Y., and Zhen, W. (2014). The horse bones unearthed from Zaoshugounao site in Chunhua County of Shaanxi Province. Journal of Northwest University 44, 311–317.

26. Bei, C. (2007). The discrimination of the ban of horses exportation in West Han Dynasty. Yinzhan Acad. J 20, 46–50.

27. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120.

28. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

29. McCormick, R.F., Truong, S.K., and Mullet, J.E. (2015). RIG: recalibration and interrelation of genomic sequence data with the GATK. G3 (Bethesda) 5, 655–665.

30. Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80–82.
50. Akey, J.M., Ruhe, A.L., Akey, D.T., Wong, A.K., Connelly, C.F., Madeoy, J., Nicholas, T.J., and Neff, M.W. (2010). Tracking footprints of artificial selection in the dog genome. Proc. Natl. Acad. Sci. USA 107, 1160–1165.

51. China National Commission of Animal Genetic Resources (2011). Animal genetic resources in China: horse, donkey and camels (China Agriculture Press), pp. 96–110.

52. Pu, Y., Zhang, Y., Zhang, T., Han, J., Ma, Y., and Liu, X. (2020). Identification of novel IncRNAs differentially expressed in placentas of Chinese Ningqiang pony and Yili horse breeds. Animals (Basel) 10, 119.

53. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578.
## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Biological samples |        |            |
| Falabella           | this study | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780752/ |
| Jinjiang horse      | this study | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780752/ |
| (CSTR:12156.05.1311C0001000001045) |        |            |
| Debao pony          | this study | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780752/ |
| (CSTR:12156.05.1311C0001000001081) |        |            |
| Ningqiang           | this study | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780752/ |
| (CSTR:12156.05.1311C0001000001050) |        |            |
| Yili horse          | this study | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780752/ |
| (CSTR:12156.05.1311C0001000001053) |        |            |
| Niqi9916            | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Deba2               | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Yili2               | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Jianchang horse     | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Chaidamu horse      | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Hequ horse          | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Datong horse        | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Jiangzi horse       | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Langkazi            | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Mozhu horse         | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Nimu horse          | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Yanqi horse         | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Erlunchun           | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Mongolian horse     | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Mongolian horse     | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Przewalskii         | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Thoroughbred        | 8 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNAS15160/ |
| Debao pony          | 7 | https://www.ncbi.nlm.nih.gov/bioproject/PRJNA641243/ |

(Continued on next page)
**Chemicals, peptides, and recombinant proteins**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| RPMI 1640           | GIBCO  | RRID: 11875093 |
| Fetal Bovine Serum  | GIBCO  | RRID: A3160501 |
| penicillin,streptomycin and L-glutamine | GIBCO | RRID: 15140122 |
| DMEM                | GIBCO  | RRID: 11-995-040 |
| BamHI restriction enzymes | NEW ENGLAND BioLabs | RRID: ER0055 |
| Nhel restriction enzymes | NEW ENGLAND BioLabs | RRID: R0131 |
| PGL4.23 vector      | Promega| RRID: 9PHE841  |
| EasyGeno            | QIAGEN | RRID: V1201  |
| pGL4.74 vector      | Promega| RRID: 9PHE692  |
| Lipofectamine3000   | ThermoFisher Scientific | RRID: L3000001 |
| Dual-Luciferase Reporter Assay System | Promega | RRID: E1910 |
| TRIzol             | Invitrogen | RRID: 15596026 |
| TIANGEN Reagent kit | TIANGEN | RRID: DP341 |
| PrimeScript RT Reagent Kit | TaKaRa | RRID: RR037A |
| SYBR Premix Ex Taq kit | TaKaRa | RRID: RR820A |
| RNA Later           | ThermoFisher | RRID: AM7020 |

**Deposited data**

| Raw and analyzed data | This study | NCBI: PRJNA780752 |

**Experimental models: Cell lines**

| Human: Human embryonic kidney T cell line: HEK293T | Beijing university medical college hospital | RRID: 4201PAT-CCTCC01347 |

**Experimental models: Organisms/strains**

| Enhancer knock-out mice (C57BL/6N) | Beijing Biocytogen | N/A |

**Software and algorithms**

| Trimmomatic | 27 | [http://www.usadellab.org/cms/?page=trimmomatic](http://www.usadellab.org/cms/?page=trimmomatic) |
| Burrows-Wheeler Algorithm (BWA) | 28 | [http://bio-bwa.sourceforge.net/](http://bio-bwa.sourceforge.net/) |
| Genome Analysis Toolkit (GATK) | 29 | [https://gatk.broadinstitute.org/hc/en-us](https://gatk.broadinstitute.org/hc/en-us) |
| SNPEff | 30 | [http://pcingola.github.io/SnpEff/se_introduction/](http://pcingola.github.io/SnpEff/se_introduction/) |
| PLINK | 31 | [https://www.cog-genomics.org/plink2](https://www.cog-genomics.org/plink2) |
| GCTA | 32 | [http://cns-genomics.com/software/gcta/mlmassoc.html](http://cns-genomics.com/software/gcta/mlmassoc.html) |
| ADMIXTURE | 33 | [http://dalexander.github.io/admixture/index.html](http://dalexander.github.io/admixture/index.html) |
| PHYLIP | 34 | [http://evolution.genetics.washington.edu/phylip.html](http://evolution.genetics.washington.edu/phylip.html) |
| Vcftools | 35 | [http://vcftools.sourceforge.net](http://vcftools.sourceforge.net) |
| Selscan | 36 | [https://github.com/szpiech/selscan](https://github.com/szpiech/selscan) |
| BioMart | | [http://www.biomart.org/](http://www.biomart.org/) |
| gProfiler | 9 | [https://bit.cs.ut.ee/gprofiler/](https://bit.cs.ut.ee/gprofiler/) |
| PopLDdecay | 37 | [https://github.com/BGI-shenzhen/PopLDdecay](https://github.com/BGI-shenzhen/PopLDdecay) |
| LDBlockShow | 38 | [https://github.com/BGI-shenzhen/LDBlockShow](https://github.com/BGI-shenzhen/LDBlockShow) |
| PALEOMIX pipeline37 | 39 | [https://github.com/MikkelSchubert/paleomix](https://github.com/MikkelSchubert/paleomix) |
| mapDamage219 | 21 | [https://ginolhac.github.io/mapDamage](https://ginolhac.github.io/mapDamage) |
| NGS QC Toolkit | 40 | [https://github.com/mjain-lab/NGSQCToolkit](https://github.com/mjain-lab/NGSQCToolkit) |
Lead contact
Further information and requests for resources, material and reagents should be addressed and will be fulfilled by the lead contact, Lin Jiang (jianglin@caas.cn).

Materials availability
All plasmids and mouse lines generated in this study are available upon request.

Data and code availability
The NCBI accession number for the genomic data reported in this paper is PRJNA780752. All other previously published genomic data used in this study are available in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Candidate SNP genotyping in a larger population
We used KASP (Kompetitive Allele-Specific PCR)43 as a cost-effective approach to genotype a panel of 763 horses and ponies originating from 26 Chinese populations at the for two TBX3 enhancer positions (TBX3-EN1 and TBX3-EN2; Data S3A and S3B). The allele-specific primers were designed by carrying the standard FAM and HEX tails and with the targeted SNP at the 30 end (Data S3C). PCR cycling was performed using the following protocol: 1) hot start, 95°C, 15 min; 2) ten touchdown cycles, 95°C, 20 s; 3) touchdown at 65°C initially and decreasing by −1°C per cycle for 25 s; and 4) annealing, 95°C, 10 s, 57°C, 60 s, 30 cycles. The genotyping of KASP markers was conducted at the China Golden Marker (Beijing) Biotech.

Dual-luciferase expression assays
Dual-luciferase assays were carried out in HEK293T (Human embryonic kidney 293 cells). HEK293T cell line was purchased from Beijing university medical college hospital. HEK293T cells were cultivated in Dulbecco’s Modified Eagle Medium DMEM basal medium (Gibco, USA), supplemented with 10% heat-inactivated fetal bovine serum and penicillin (0.2 U/mL)/streptomycin (0.2 μg/mL)/L-glutamine (0.2 μg/mL) (GIBCO). The 63-bp TBX3 enhancer region carrying both TBX3-EN1 and TBX3-EN2 mutations was isolated by PCR with the primers listed in Data S3D, following purification and digested with KpnI and NheI restriction enzymes before being inserted into PGL4.23-Basic vectors (EasyGeno, QIAGEN). A total of four vectors displaying different combinations of mutations of TBX3-EN1 and TBX3-EN2 were constructed and tested with reporter assays for regulatory expression effects. These included: the horse derived mutation (T and G), the TBX3-EN2 single mutation (C and G), TBX3-EN1 single mutation (T and A), and the double TBX3-EN1 and TBX3-EN2 ancestral allele (C and A), where the first letter provides the TBX3-EN1 allele, and the second letter the TBX3-EN2 allele. The TBX3 enhancer region was constructed into dual luciferase reporter vector pGL4.74 and transfected into HEK293T cell by Lipofectamine3000 (ThermoFisher Scientific, USA), following the procedure provided by the manufacturer. The transfected cells were lysed by the Dual-Luciferase Reporter Assay System (Promega E1910) and the luciferase intensity was detected using the Tecan Infinite 2000 Pro instrument. Both Firefly and Renilla luciferase activities were sequentially measured in each individual sample and expression values were calculated as the ratio of Firefly to Renilla luciferase activities (Fluc/Rluc).

Knock-out mice preparation
TBX3 enhancer−/− mice (C57BL/6N) were generated by CRISPR/Cas9 technology at Beijing Biocytogen, and housed in the pathogen-free facilities of China Agriculture University. They were resulted from heterozygous crosses between TBX3 enhancer+/− males and TBX3 enhancer−/− females. Animal experiments were performed in accordance with the regulations and guidelines established by the Animal Care Committee of the Institute of Animal Science of Chinese Academy of Agricultural Sciences (ethical permits IAS2019-24).

A total of 13 sgRNAs were first designed using the CRISPR design tool (https://wge.stemcell.sanger.ac.uk/) and the following sequence was used as input (GGGTCTTGGGAGGTCAGTC[A/G]TAATTGGCGGAAGTTT). The sgRNAs were connected to the pCS-3G vector by Gibson connection, and recombinant plasmids were validated using the Sanger sequencing (GenScript Biogene company, Nanjing, China). UCA CRISPR/Cas9 rapid construction and activity detection kit (Biocytogen, Beijing, China) was used to evaluate targeted cleavage ability of the sgRNAs. Validated Cas9/sgRNA plasmids were transfected into HEK293T cells and

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| HISAT               | 41     | http://www.ccb.jhu.edu/software/hisat/index.shtml |
| DESeq              | 42     | http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html |

REAGENT or RESOURCE SOURCE IDENTIFIER

HISAT 41 http://www.ccb.jhu.edu/software/hisat/index.shtml
DESeq 42 http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html
measured for Luciferase reporter activity relative to a control group consisted of non-transfected HEK293T cells. The sgRNA showing the highest reporter activity was selected for ligation into the T7 promoter plasmid vector, and then transcribed in vitro to obtain RNA suitable for microinjection into the mouse zygotes.

Four-week-old C57BL/6N female mice were selected as receptors. Super-ovulating female mice were obtained by first injecting 5 IU pregnant mare serum gonadotropin (PMSG) followed by 5 IU human chorionic gonadotropin (hCG) 48 h later. Super-ovulating mice were then crossed with male mice and the Fallopian tubes of resulting 1.5-day-old gestational mice were dissected and placed in glass dishes containing M2 medium. Finally, 2-cell stage embryos were harvested by using a syringe with a blunt 4" needle and rinsed the oviduct part with M2 medium.

Cas9/sgRNA and targeting vector were microinjected into the mouse zygotes and placed in a 37°C, 5% CO₂ incubator for 1.5 days. Female mice with clear signs of estrus were used as recipient surrogates. Two-cell stage embryos were selected and injected into the uterus of the surrogate mother. All resulting mice were delivered by natural birth. PCR and Southern blotting were used to genotyping of KO mice. The heterozygous TBX3 enhancer-KO mice were used for breeding a population of sufficient size for measuring body and limb bone sizes.

**RT-qPCR experiments**

We collected tissues including skin, muscle and limb bones during both the embryonic period and the 2-weeks developmental stage. All the tissues were treated by the TRIzol (Invitrogen, Massachusetts, USA) and the TIANGEN Reagent kit (TIANGEN, Beijing, China) was used to extract total RNA. Importantly, limb bones had to be powdered following grinding in liquid nitrogen before the TRIzol treatment. Total RNA was reverse-transcribed into complementary DNA (cDNA) using the PrimeScript RT Reagent Kit (TaKaRa, Dalian, China). TBX3 (forward primer: CCACCGGTTCTCT- CAATTTGACAG; reverse primer: CGGAAGCCATTGTGGTTAAAGCTG) expression levels were measured through quantitative real-time PCR on ABI 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) using the SYBR Premix Ex Taq kit (TaKaRa) and normalized to a β-actin (forward primer: CATCGATAGACCTCT ATGCCAC; reverse primer: ATGGAGCAAGCCGCATCCAC). Three replicates were performed for each sample and the average value was used for further analysis. Fold expression changes were determined using a standard 2-ΔΔCT method that compares the CT (cycle threshold) values of a reference gene (here, β-Actin) to the gene of interest for the ΔΔCT calculation and compares the ΔCT value of a reference sample with the sample of interest for the ΔΔCT calculation.

**Bone length measurements**

Littermate mice were used in all our experiments. The length of limb bones was measured on 2-week old mice and analyzed by the Micro-CT and RadiAnt DICOM viewer (5.0.1) software. The length of limb bones from 11-week old mice was measured and analyzed by X-ray and MicroDicom viewer (3.2.7) software. Specific measurement indicators are detailed in Figure 3D. Statistical significance of bone measurements was assessed using two-tailed Student’s t tests.

**METHOD DETAILS**

**Pony and horse samples collection and WGS sequencing**

We sequenced 57 novel pony and horse genomes at an average depth-of-coverage of 11.1X (Data S1A). To achieve this, we collected a minimum of two separate herds for each breed and/or location, disregarding genetically related individuals up to three-generations. Peripheral venous blood was obtained from jugular vein and rapidly frozen to −20°C. Genomic DNA was extracted from the blood using the standard Promega extraction (Promega, A1125, America). The quality and integrity of the extracted DNA was examined by the A260/A280 ratio and agarose gel electrophoresis. Genomic libraries with insert size around 350 bp were constructed following the Illumina standard protocol (Illumina Inc.) and sequenced by the Hiseq XTen platform (2x150 paired-end mode). Sequencing, sequence cleaning and alignment against the EquCab3 horse reference genome were carried out by the BerryGenomics Company (Beijing, China), following the methodology described early. Briefly, low quality raw sequence reads were filtered and the remaining reads were trimmed for Illumina adapters using the Trimmomatic-0.33. Sequence alignment was carried out using the Burrows-Wheeler Algorithm (BWA) implemented in the BWA-mem module with default parameters. Paired reads that were mapped to the exact same position on the reference genome were removed with the MarkDuplicates in Picards (picard-tools-1.56 at http://picard.sourceforge.net). Additional realignment around indels was performed by the Genome Analysis Toolkit (GATK). The sequence data obtained were sufficient to characterize individual genomes at minimum 11.1X depth-of-coverage and 96.3% genome coverage, which allowed for the identification of 20,428,318 SNPs with high confidence (Data S1B), following the early methodology from Hwang et al. SNP variation was identified in each individual sample using the HaplotypeCaller (with the following parameters:--pairHMM VECTOR LOGLESS CACHING--emitRefConfidence GVCF--variant_index_parameter 128000) in the GATK (version 4.0). SNP calling of multiple samples was then performed using the GATK GenotypeGVCFs with default settings. SNPs that did not meet the following criteria were excluded: (1) the mean sequencing depth over all included individuals was superior to 3 x or inferior to 30 x; (2) the minor allele frequency was superior to 0.05; (3) the missing rate was inferior to 0.1; and (4) the locus was biallelic. All SNPs were annotated using the SNPEff 4.0 based on the gene annotation of the reference genome provided by the NCBI.

The 57 genomes sequenced in this study were complemented with a previously released data consisting of 113 individual horse genomes and 17 Debao ponies, providing a panel of 187 genomes. The ancient genome time-series was obtained from Fages and
collaborators and included genome-wide sequence data from 245 equine subfossils from across Eurasia and spanning the last six millennia. Modern horses were binned into two main groups, consisting of the ‘Pony’ and the ‘Horse’ group. The former included DeBao pony (DeBa, N = 27), NingQiang pony (NiQi, N = 8), JiangChang pony (JiCh, N = 8), JinJiang pony (JiJi, N = 9) and Falabella pony (FLBL, N = 23). The ‘Horse’ group included the following breeds: ChaiDamu (CDM, N = 10), HeQu (HeQu, N = 6), DaTong (DaTo, N = 10), JiangZi (JiZi, N = 10), LangKazi (LKZ, N = 10), NiMu (NiMu, N = 9), MoZhongka (MoZh, N = 6), Inner Mongolia (MoGo, N = 8), Mongolia (WMG, N = 7), ErLunchun (ELC, N = 7), YiLi (YiLi, N = 10), YanQi (YaQi, N = 7), Thoroughbred, (ThB, N = 7), and the Przewalski’s horse (PrZ, N = 5). Full sample details are presented in key resources table.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Population structure**

Population structure among all modern horse and pony breeds was investigated using a total of 11,553,176 high-quality SNPs, following conversion of the VCF files into the PLINK format by the VCFTools, and filtering for minor allele frequency (MAF) > 0.05, call out rate > 0.9 and hwe > 1e-6. Pruning was carried out within windows of 1,000 SNPs (step-size 5 SNPs), using the –ind-pairwise 1000 5 0.5 parameters in the PLINK. This provided a total of 1,593,119 unlinked SNPs for PCA, ADMIXTURE and phylogenetic reconstruction. The PCA was conducted using the GCTA software, the first two axes were plotted using with the R program (version 3.5.0). Individual ancestry profiles were inferred using 10,000 bootstrap replicates in the ADMIXTURE (version 1.3.0) and assuming 2 to 4 ancestral populations (K). Neighbor-joining tree was constructed using the PHYLIp v3.68 based on the pairwise genetic distance matrix returned by the PLINKv1.9.

**Runs of homozygosity and linkage disequilibrium**

Runs of homozygosity (ROH) were identified for each breed/population applying the ‘runs of homozygosity’ function in the program PLINK v1.0. The unpruned matrix of high-quality variants. The analysis was run with the following options ‘–homozyg-window-kb 1000–homozyg-window-snp 50–homozyg-window-het 1–homozyg-snp 10–homozyg-kb 100–homozyg-density 10–homozyg-gap 100’. Patterns of linkage disequilibrium (LD) among different horse breeds were reconstructed from individual VCF files using the PopLDDecay with default parameters, except that the maximum distance between two SNPs was set to 500 kb.

**Selection scans**

We used all SNPs passing the quality control to detect signatures of selection in Chinese ponies and Falabella ponies within the panel of 187 genomes. Since Falabella ponies were imported from France (N = 23), the analyses contrasting selection signatures in horses and ponies were repeated in Falabella and Chinese ponies, separately. Genome-wide distribution of FST fixation indices between Chinese ponies and horses were calculated following. Chinese ponies included 52 samples from four breeds (Deba = 27, JiCh = 8, NiQi = 8 and JiJi = 9), while Chinese horses included 100 individuals across 12 breeds (CDM = 10, HeQu = 6, DaTo = 10, JiZi = 10, NiMu = 9, MoZh = 8, LKZ = 10, MoGo = 8, WMG = 7, ELC = 7, YiLi = 10 and YaQi = 7). FST fixation, XP-EHH and nucleotide diversity ratio, inbreeding coefficient (FIS), were calculated within 100 kb sliding windows, considering a 15 kb overlap between adjacent windows. Those windows comprised within the top-1% quantile of all three statistics were considered as candidate selection targets and annotated using the genomic database search engine BioMart (Data S2A–S2D). The selection detection analysis of FST fixation was conducted in Falabella ponies versus Chinese horses and Thoroughbred horses. Functional enrichment for Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses was tested with goProfiler. The enrichment significance was assessed for g:SCS ‘Set Counts and Sizes’, which is more conservative than Benjamini-Hochberg False Discovery Rate but not as strict as Bonferroni correction. To investigate whether different genetic pathways may underpin the short stature of Chinese ponies and Western ponies, we downloaded whole genome sequencing data of 24 ponies and 40 horses from the European Nucleotide Archive (https://www.ebi.ac.uk) and SRA (https://www.ncbi.nlm.nih.gov/sra/) (Data S2). We then combined these data with FLBL ponies sequenced in this study to form a group of ‘Western pony breeds’ including FLBL (N = 23), Shetland pony (N = 18), German Riding pony (N = 2), Haflinger (N = 3) and Icelandic pony (N = 1), as well as a group of horses, including Akhal-Tekе (N = 3), Arabian (N = 3), Hanoverian (N = 5), Holsteiner (N = 5), Oldenburg (N = 3), Trakehner (N = 2), Franches-Montagnes (N = 15), Morgan (N = 1), Swiss Warmblood (N = 1), Dutch warmblood (N = 1) and Standardbred (N = 1). The methodology then followed the same analytical steps as described for the Selection scan carried out between Chinese ponies and Chinese horses. The selection candidates identified among ‘Western pony breeds’ and the horse group are provided in Data S2J and S2K.

**Genome wide association study**

The wither height of horses was estimated from the data provided by the China National Commission of Animal Genetic Resources following random sampling from normal distributions centered around the breed mean wither height within standard error, and using 100 pseudo-replicates. Association analysis was performed by the PLINK version 1.9 (~association), applying a significance criterion of p < 6e-8 and considering a genomic correction normalized by the effective number of SNPs considered. The first five PCA components were used for accounting population structure (~adjust-covar). Additionally, we constructed haplotype blocks (~blocks) using the PLINK v1.07, together with the ‘~hap plink.blocks ~hap-assoc’ and ‘~hap-freq’ options, allowing for the discovery of haplotype associations and frequency estimates in each block. We focused on SNPs: 1) present within statistically significant
blocks (p < 0.05); 2) showing different haplotype frequencies; and 3) higher than GWAS threshold of 6e-8. The LDBlockShow was then used to visualize the linkage disequilibrium (LD) blocks within 200 kb in length. 38

Ancient allelic trajectories
Sequencing data were downloaded from the European Nucleotide Archive and processed using the PALEOMIX pipeline39 to retrieve individual BAM alignments against the EquCab3 horse reference genome, following the methodology from Fages and collaborators. 19 Allelic trajectories were assessed by grouping individuals within time bins of 1,000 years (step-size 500 years), and randomly sampling one read per position covered in 1,000 pseudo-replicates.

Post-mortem DNA damage parameters were inferred from individually trimmed and rescaled BAM alignments in the mapDamage2.21 Reads counts were calculated with the ANGSD (-doCounts 1 -dumpCounts 4) and genotyped by in-house python with random genotyping at sequencing coverage more than three.

Transcriptome sequencing and analysis
A total of seven placenta samples were randomly taken from the Chinese NiQi ponies and YiLi horses originating from the Shaanxi and Xinjiang provinces.52 These samples were genotyped using the Sanger sequencing of PCR products as AA and GG homozygotes at the TBX3-EN2 locus, respectively. Tissues samples were placed in RNA-later directly after collection. Total RNA was extracted using the RNeasy Mini Kit (QIAGEN) and quantified on the Agilent 2100 Bioanalyzer device (Agilent technologies, Santa Clara, CA, USA). RNA-Seq libraries were constructed using the mRNA-Seq sample preparation kit (Illumina) according to the manufacturer’s protocol and sequenced on the Illumina XTen platform (2x150 paired-end mode, Illumina, by the BerryGenomics Company (Beijing, China)). After quality control by the NGS QC Toolkit v2.3.3 software, we followed the methodology from Pertea and colleague40 to characterize differences in the horse and pony transcriptomes. First, the HISAT software (Hierarchical Indexing for Spliced Alignment of Transcripts) 41 was used with default settings to map reads against the reference horse genome (EquCab 3.0).45 Then, the StringTie was used to assemble alignments into full transcripts. Finally, expression outputs from the StringTie were processed applying the rigorous statistical methods implemented in the Ballgown software to identify differentially expressed transcripts and genes between Chinese NiQi ponies and YiLi horses. FPKM (fragments per kilobase of exon per million fragments mapped) expression values were calculated using a combination of R packages, including the ballgown, RSkittleBrewer, genefilter, dplyr and devtools, while the cummberbund R package was used for gene expression data visualization.53 FPKM values for all the annotated transcripts were processed through PCA using ‘gmodels’. The differentially-expressed genes (DEGs) were defined by fold changes superior or equal to 2 (p value ≤ 0.05) in the DESeq 42 (Data S2N).