Diagram

Description automatically generated

**Figure S1.** The hypothetical genetic distribution of liabilities across genetic risk groups. Control included the Belgian (BE), Icelandic Horse (IH), Shetland Pony (SP) and Welsh Pony (WP) breeds; low risk included the Lusitano (LU) and Arabian (AR) breeds; medium risk included the Standardbred (SB), Thoroughbred (TB), Quarter Horse (QH) breeds, and high risk consisted of the Peruvian Horse (PH). The bottom plot showed distributions of liabilities for each risk group separately estimated by a Bayesian linear mixed model. The numbers are mean (standard deviation) for each distribution.

**Shape

Description automatically generated**

**Figure S2**. Genomic relationship matrix heatmap plotted for each breed of horse in the study. The colors represent the strength of the relationship between individual horses within a breed.



**Figure S3**. The plot shows the variance explained by the first ten principal components. The solid line represents the variance reduction trend over all ten components. The analysis used all SNPs included in the GWAS.

Chart, line chart

Description automatically generated

**Chart, bar chart

Description automatically generated**

**Figure S4**.Admixture output for the four lowest mean squared error (MSE) values obtained by cross validation over K = 8, 12, 14, and 18. The lowest value was at K=14. Ancestry admixture proportion within each breed is represented by different colors. In all these scenarios, the Standardbred, Peruvian Horse, Lusitano, and Arabian showed some degree of ancestral gene flow based on admixture analysis.

Chart

Description automatically generated

**Figure S5**. The decay of LD by distance in ten horse breeds plotted for marker distances up to 100 kb. The y-axis shows distribution of average pairwise r2 values.

Chart

Description automatically generated

**Figure S6**. Estimated effective population sizes (Ne) in ten horse populations over 100 generation.

Chart, line chart, histogram

Description automatically generated

**Figure S7**. The plot shows the variance explained by the first ten principal components. The solid line represents the variance reduction trend over all ten components. The analysis used only the DSLD-associated significant GWAS SNPs.