

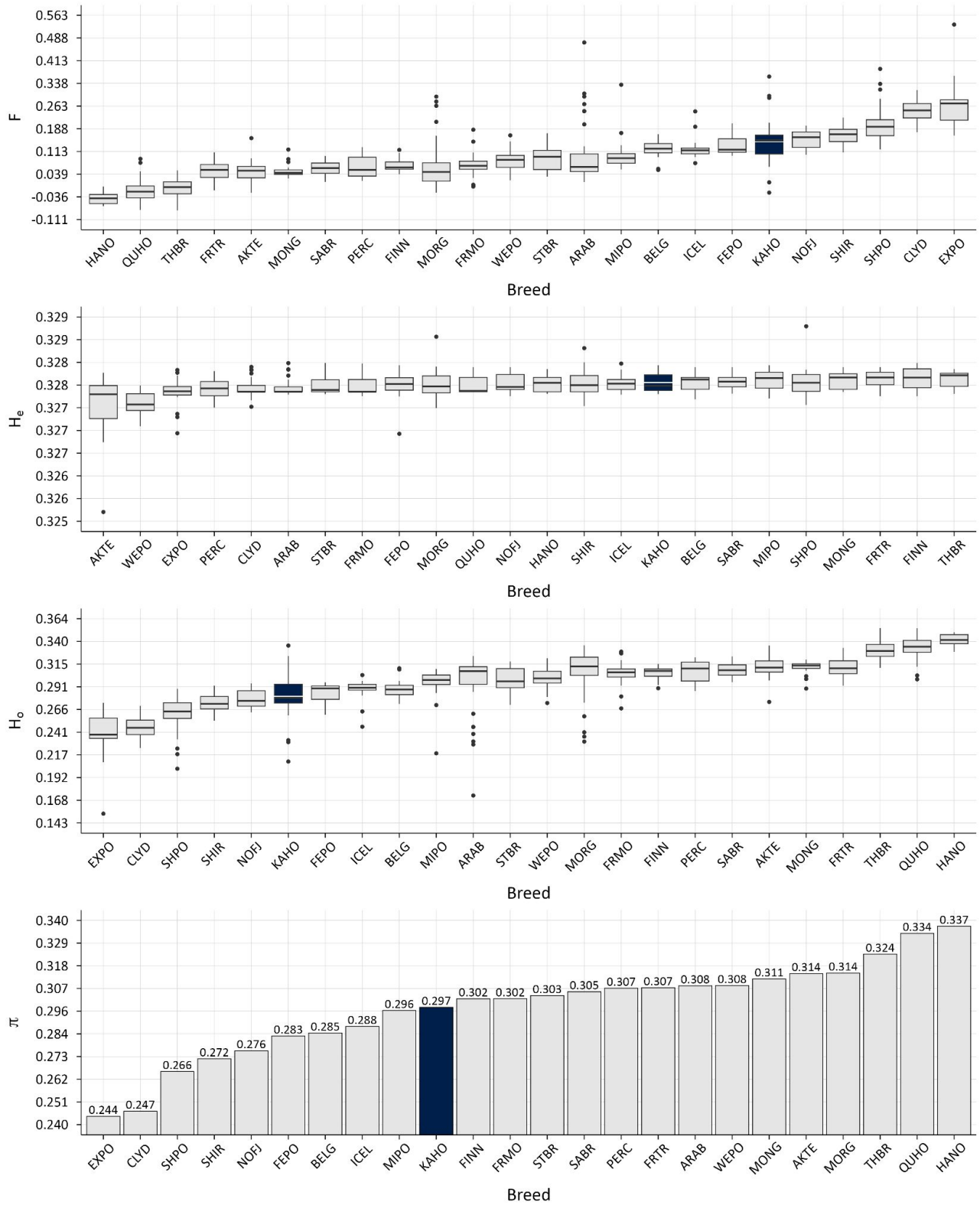


Figure S1. Inbreeding coefficient (F), expected heterozygosity (H_e), observed heterozygosity (H_o), and nucleotide diversity (π) calculated using the Combined Autosomal SNP Dataset. Kaimanawa Horses are shown in blue, and all other breeds in gray. Breed abbreviations are listed in Table 1. F , H_e , H_o were calculated using PLINK v1.9 [32, 33], and π was calculated using VCFtools v0.1.16 [45].

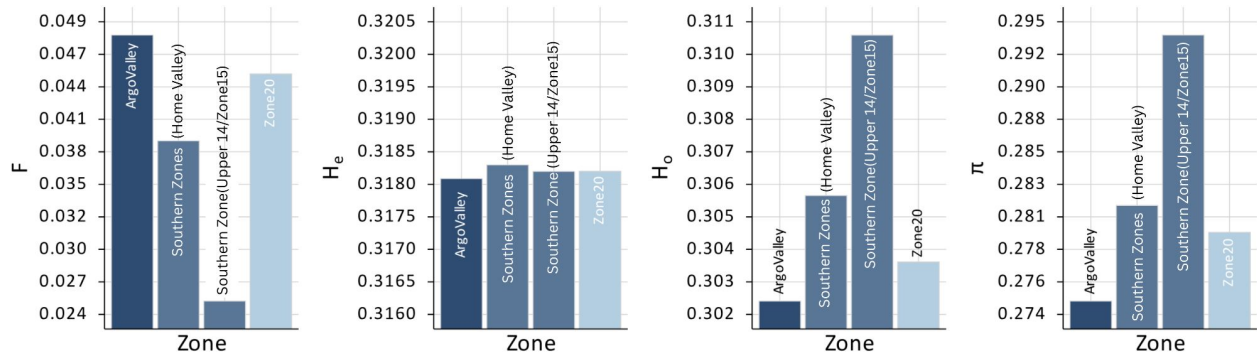


Figure S2. Genetic diversity in Kaimanawa Horses based on the Combined Autosomal SNP Dataset, shown across owner reported capture zones (Dataset 1). The reported measures include the observed individual inbreeding coefficient (F), expected heterozygosity (H_e), observed heterozygosity (H_o), and nucleotide diversity (π). F , H_e , H_o were calculated using PLINK v1.9 [32, 33], and π was calculated using VCFtools v0.1.16 [45].

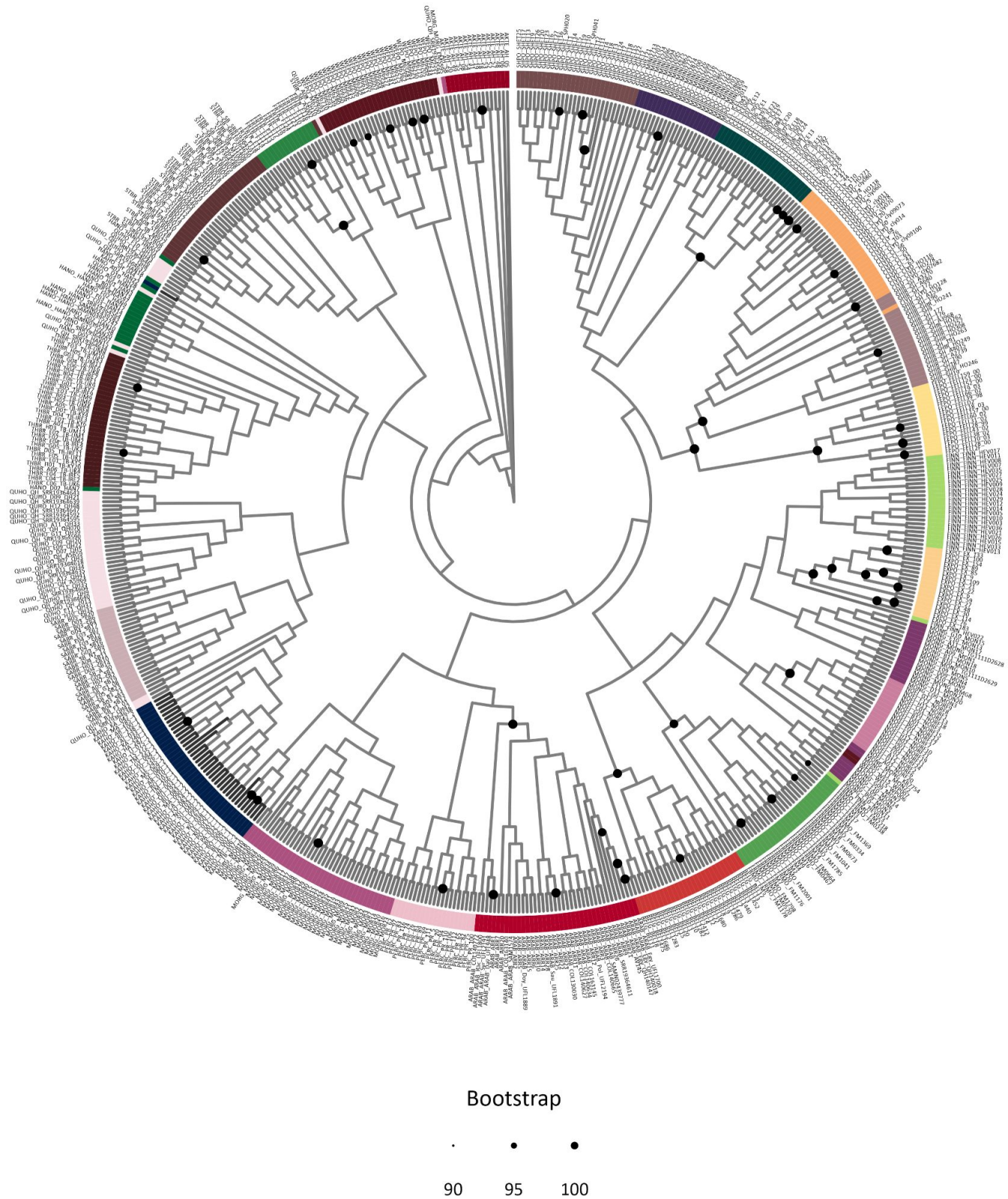


Figure S3. Circular maximum likelihood tree (ML) with equalized branch lengths based on the Combined Autosomal SNP Dataset. Bootstrap support values > 50 are indicated by standardized dot size at the corresponding branches. The outer colored ring represents breed identity . with breed abbreviations and colors as defined in Table 1. The ML tree was inferred using IQ-TREE v2.3.6 [48] and visualized in R v4.4.2 [44] with ggplot2 v4.0.0 [50] and ggtree v3.14.0 [51].

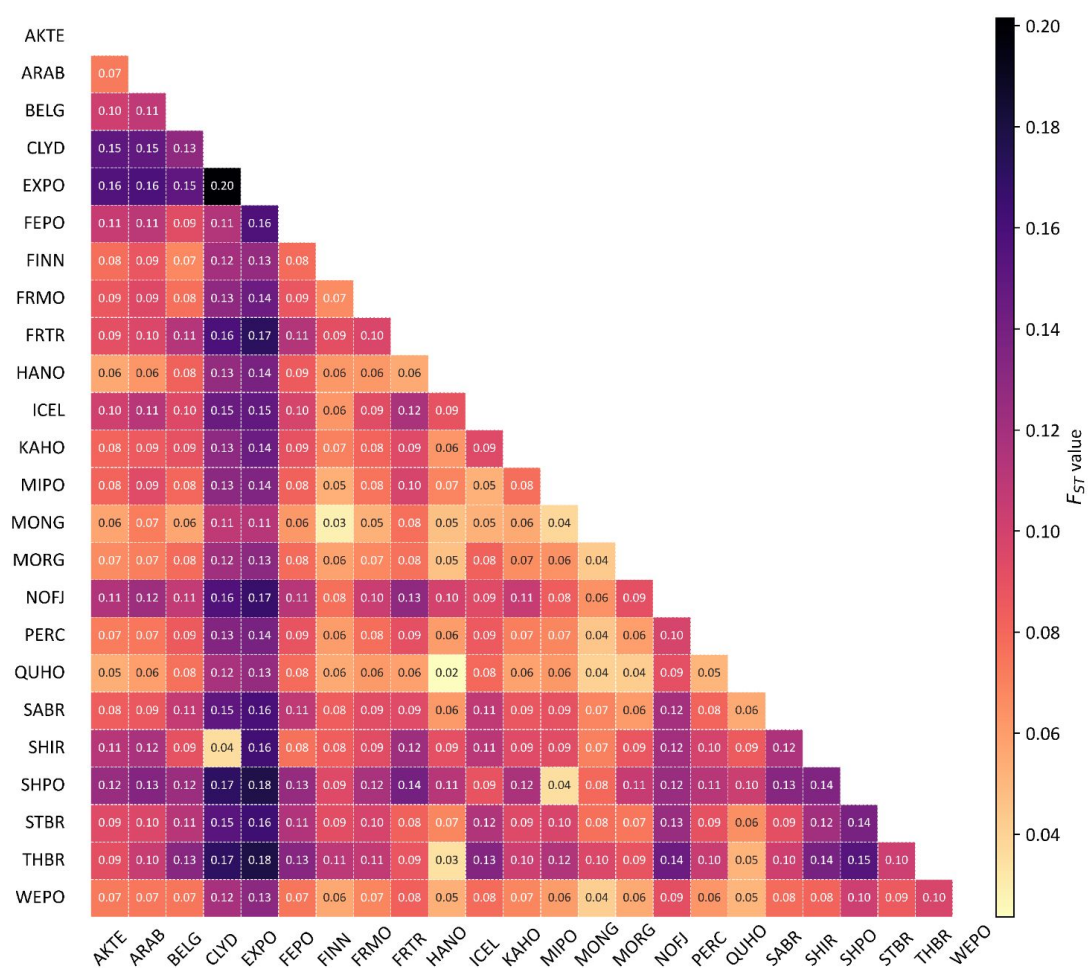


Figure S4. Pairwise F_{ST} matrix based on the Combined Autosomal SNP Dataset, calculated by 1000 repetitions and 95% confidence intervals. All comparisons exceeded the significance threshold (p -value < 0.001). Color scale: dark purple indicates the highest differentiation (up to 0.201), and yellow the lowest (down to 0.000). Comparisons with the Kaimanawa Horse population are outlined in dark blue. Pairwise F_{ST} values and p -values were calculated using the R v4.4.2 [44] package StAMPP [54]. The breed abbreviations are listed in Table 1.

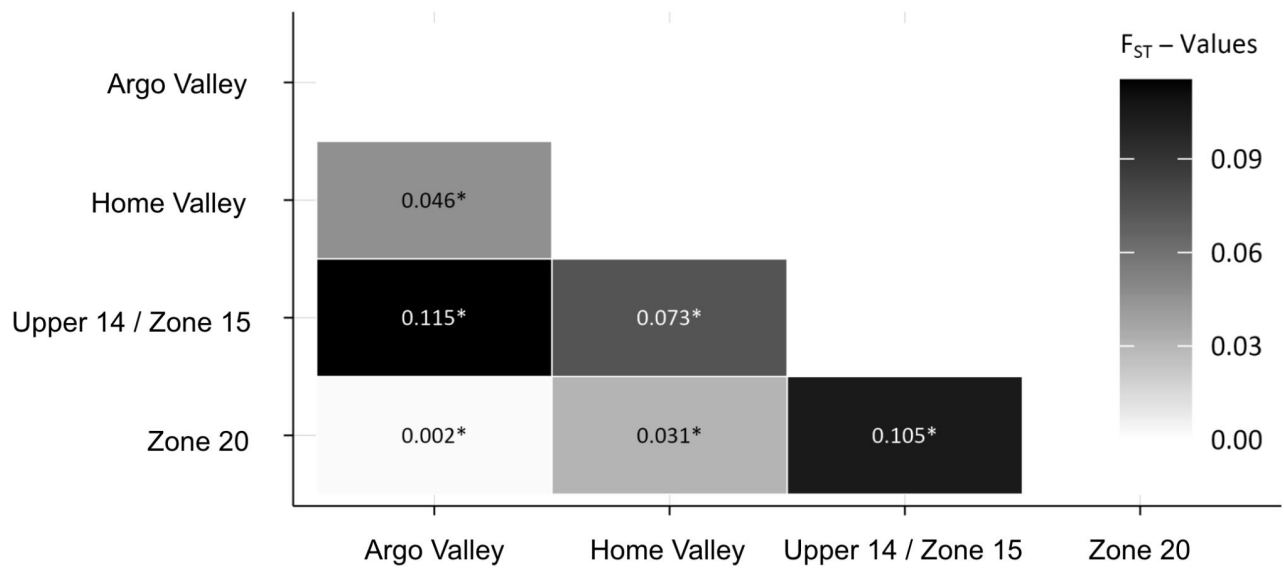


Figure S5. Pairwise F_{ST} matrix of Kaimanawa Horses from the Combined Autosomal SNP Dataset, grouped by zones (Dataset 1). Values were calculated using 1,000 repetitions with 95% confidence intervals. Comparisons with p-values above the significance threshold of 0.005 are indicated by an asterisk (*). Color scale: black indicates the highest differentiation (up to 0.115), and white the lowest (down to 0.000). Pairwise F_{ST} values and p-values were calculated using the R v4.4.2 [44] package StAMPP [54].

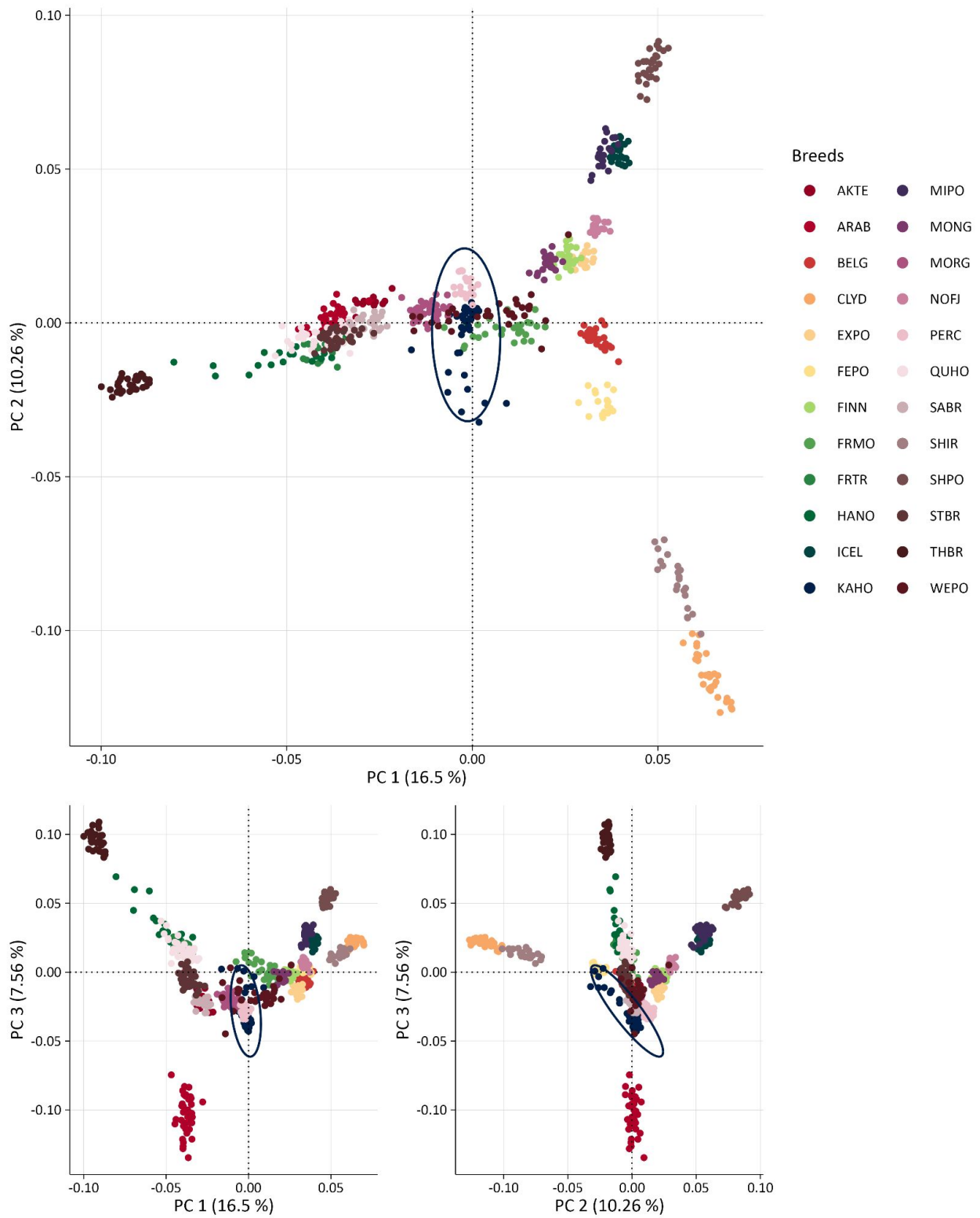


Figure S6. Principal component analysis (PCA) plot showing PC1 (16.5%) vs PC2 (10.26%), PC1 (16.5%) vs. PC3 (7.59%) as well as PC2 (10.26%) vs. PC3 (7.59%) of the total genetic variance. Each point represents one individual of the Combined Autosomal SNP Dataset colored by breeds as indicated in the legend. The blue ellipse denotes the 95% confidence interval for Kaimanawa Horses. Breed abbreviations and coloration are as listed in Table 1. PCA was calculated using PLINK 1.9 [32, 33].

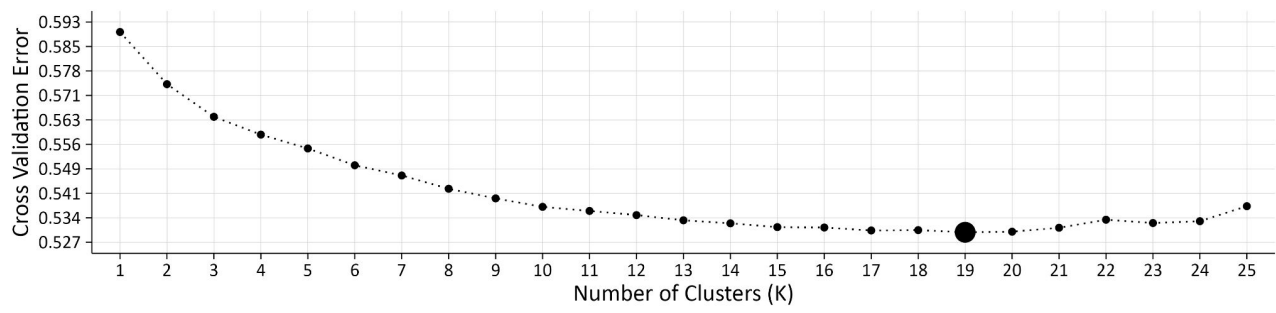


Figure S7. Admixture cross-validation (CV) error plot based on the Combined Autosomal SNP Dataset (24 breeds, 629 individuals, and 17,981 SNPs). The lowest CV error was observed at $K = 19$, indicating the optimal number of genetic components for the dataset. Admixture analysis was performed using ADMIXTURE v1.3.0 [57].

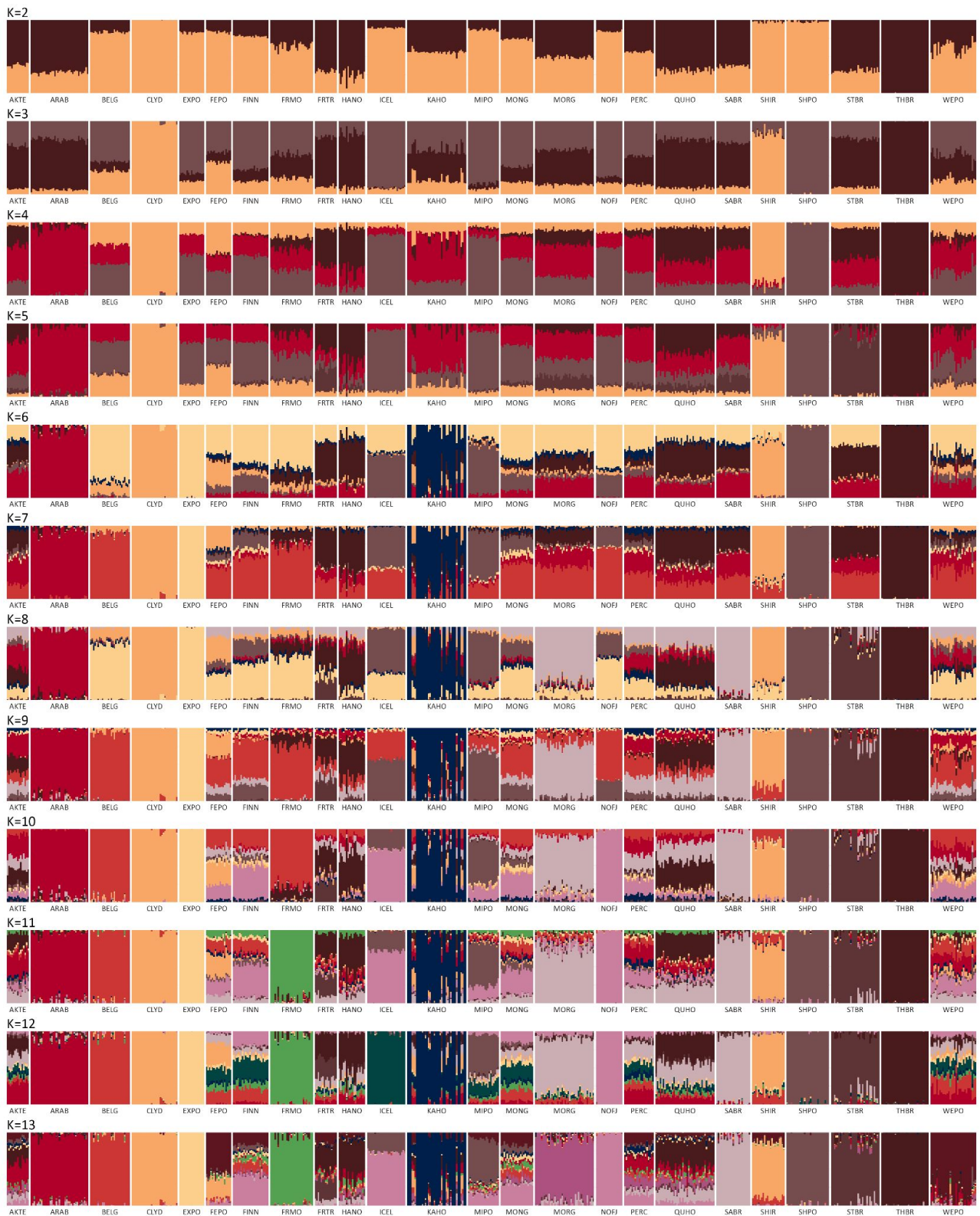




Figure S9. Admixture plots ($K = 14 - 25$) based on the Combined Autosomal SNP Dataset comprising 24 breeds, 629 individuals, and 17,981 SNPs. The color scheme for each K reflects the breed in which the respective genetic component was most prevalent. The lowest cross-validation error was observed at $K = 19$. Breed abbreviations and respective colors are as described in Table 1. Admixture analysis was performed using ADMIXTURE v1.3.0 [57].

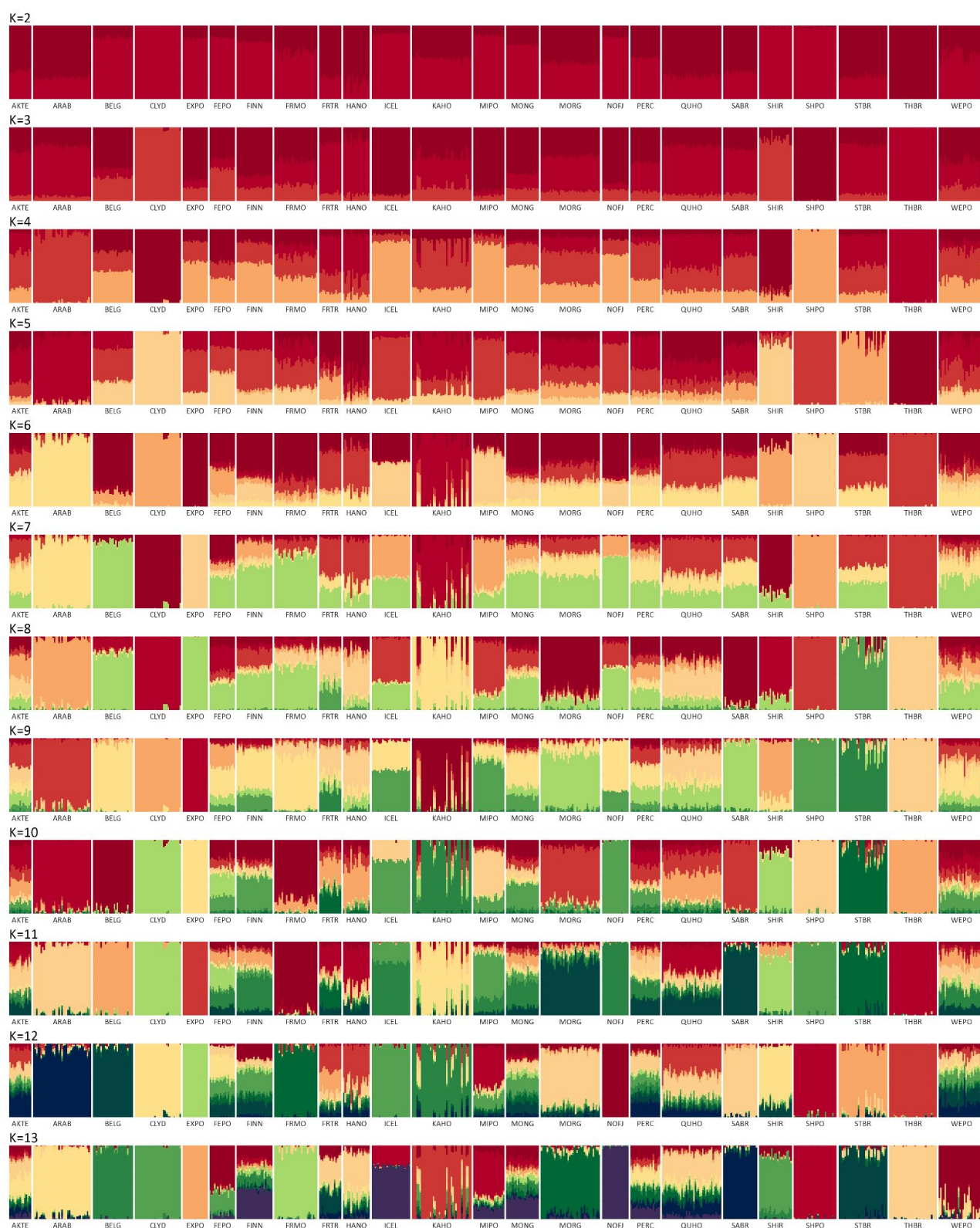


Figure S10. Admixture plots ($K = 2 - 13$) based on the Combined Autosomal SNP Dataset (24 breeds, 629 individuals, and 17,981 SNPs.) Each K is colored in increasing order. The lowest cross-validation (CV) error was observed at $K = 19$. Breed abbreviations and respective colors are listed in Table 1. Admixture analysis was performed using ADMIXTURE v1.3.0 [57].

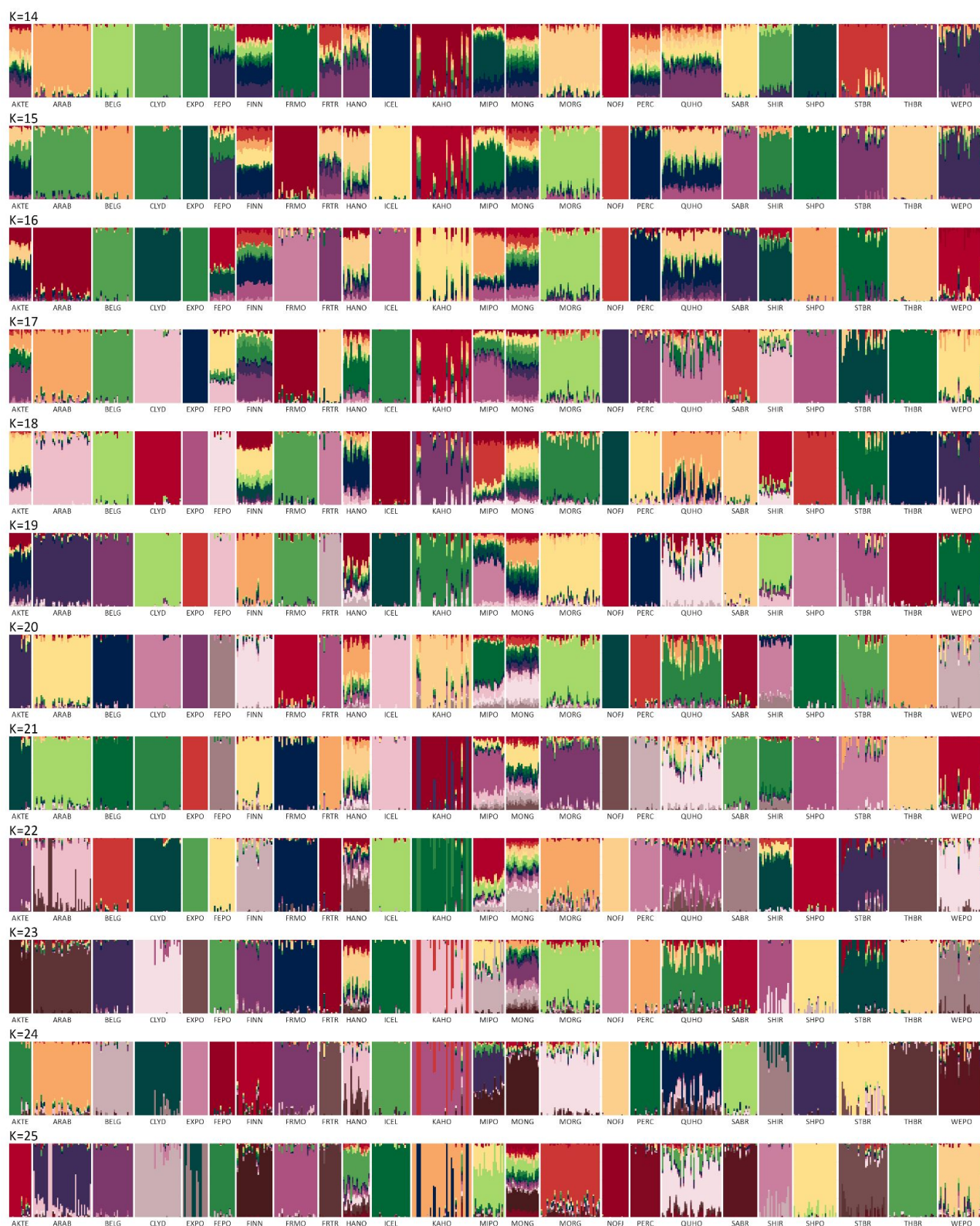


Figure S11. Admixture plots (K = 14 - 25) based on the Combined Autosomal SNP Dataset (24 breeds, 629 individuals, and 17,981 SNPs). Each K is colored in increasing order. The lowest cross-validation error was observed at K = 19. Breed abbreviations and respective colors are listed in Table 1. Admixture analysis was performed using ADMIXTURE v1.3.0 [57].

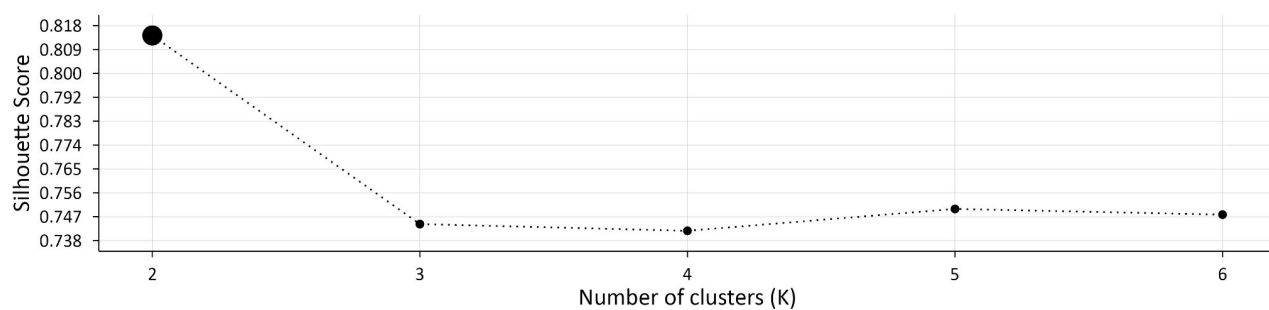


Figure S12. Silhouette score for K = 2 – 6 based on PC-AiR [58] clustering of 142 Kaimanawa Horses using 55,423 SNPs (Dataset 2). Silhouette analysis was performed in R v4.4.2 [44] survival package v3.7 [87]

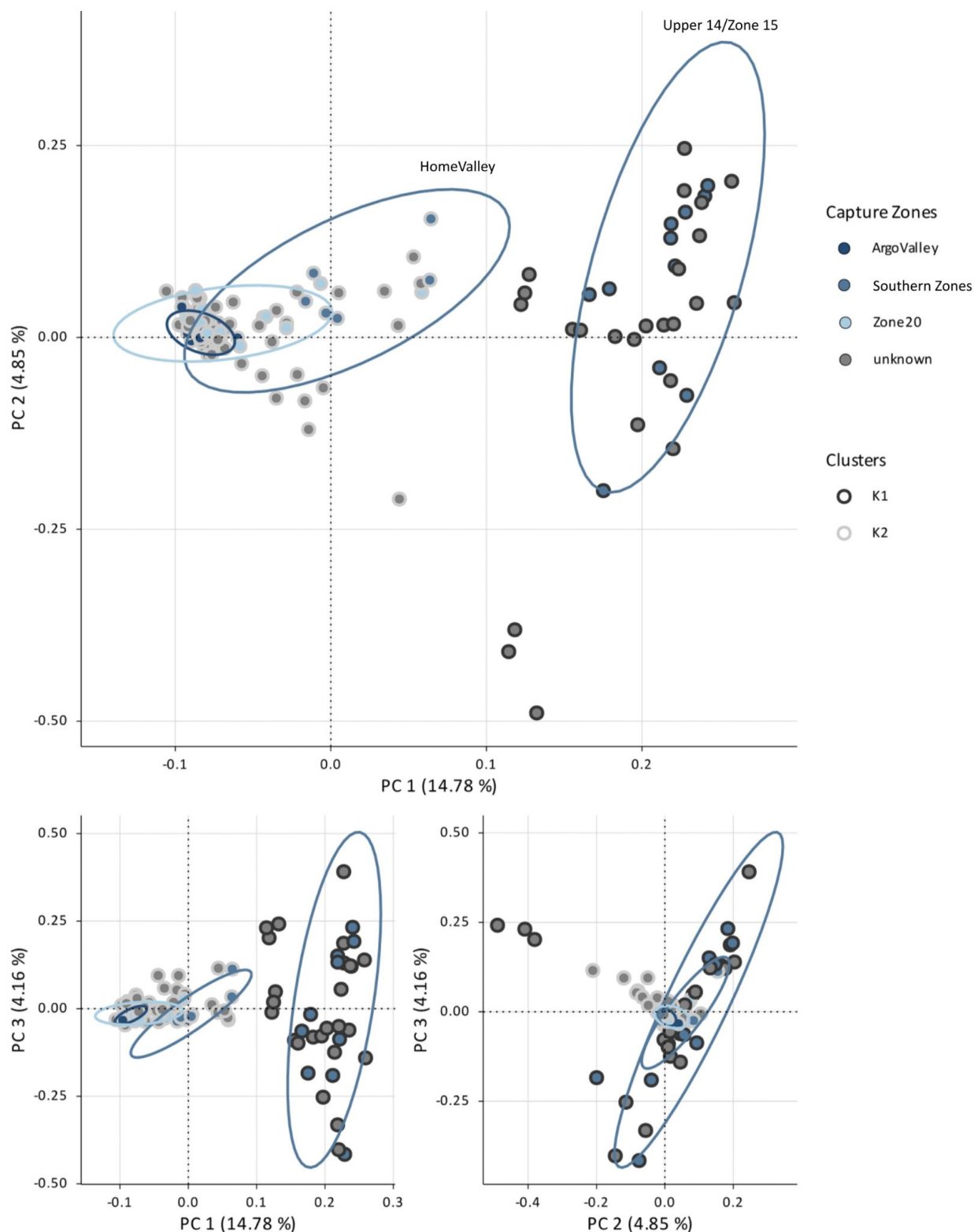


Figure S13. PC-AiR plots [58] with $K = 2$ cluster assignments of 142 Kaimanawa Horses based on 55,423 SNPs (Dataset 2), showing PC1 (15.78%) vs PC2 (4.85%), PC1 (14.78%) vs. PC3 (4.16%) as well as PC2 (4.85%) vs. PC3 (4.16%) of the total genetic variance. Horses with owner reported capture zones are shown in different shades of blue, representing areas of the Kaimanawa Ranges: Argo Valley, Southern Zones and Zone 20. Horses with unknown capture zones are shown in gray. Circles represent the 95% confidence intervals (CI) for each geographic group and are colored to match their respective groups. Circles around the data points indicate cluster assignments: light gray for $K = 1$ and dark gray for $K = 2$. PC-AiR [58] was performed in R v4.4.2 [44] using the GENESIS package v2.36.0 [60], together with SNPRelate v1.40.0 and GWASTools v1.52.0 [59].

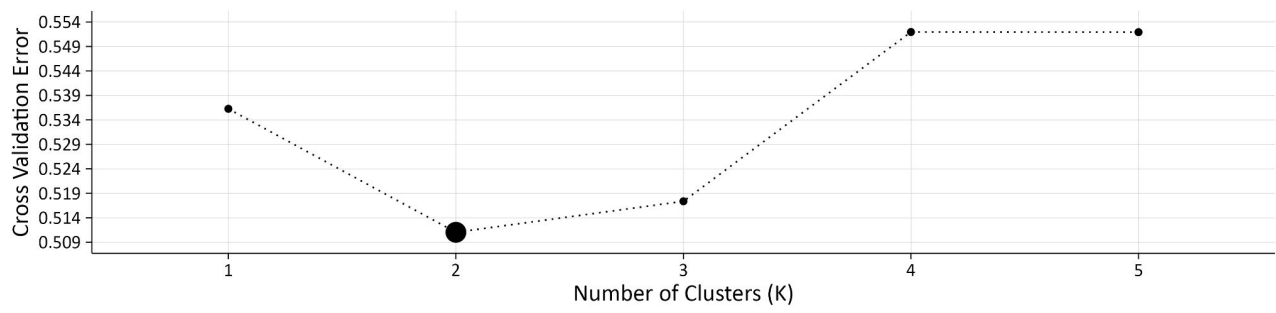


Figure S14. Admixture cross-validation (CV) error plot for Kaimanawa Horse individuals ($n = 90$) based on Dataset 3 (27,278 SNPs). CV error was calculated for $K = 2 - 5$, with the lowest value observed at $K = 2$, indicating the best-supported model for the dataset. Analyses were performed using ADMIXTURE v1.3.0 [57].

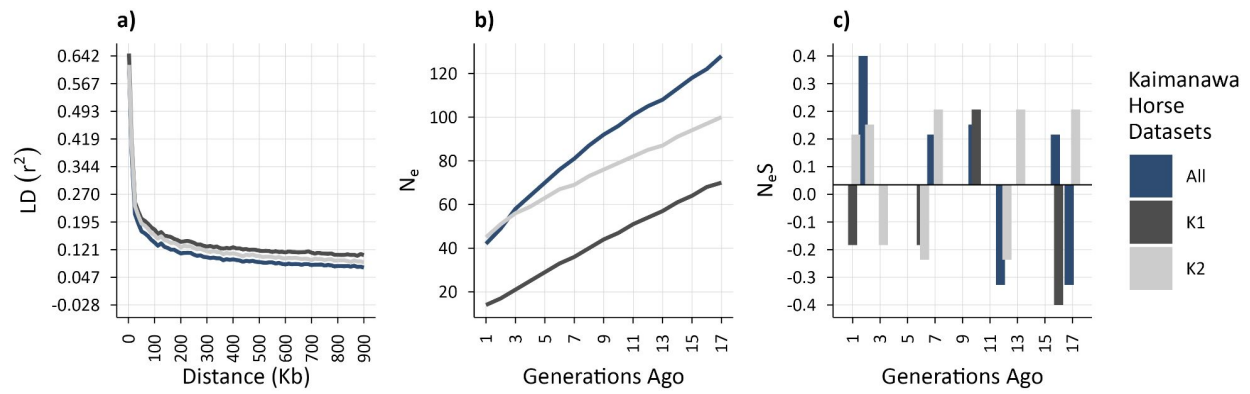


Figure S16. a) Linkage disequilibrium (LD) decay plot showing the decrease in non-random association between alleles with increasing physical distance (kb). b) Historical effective population size (N_e) of Kaimanawa Horses over the last 17 generations. c) Slope analysis ($N_e S$) of historical N_e across the same period, illustrating the rate and direction of demographic change. Positive values indicate an accelerated decline, whereas negative values indicate a slower decline. Analysis was performed using Dataset 4.

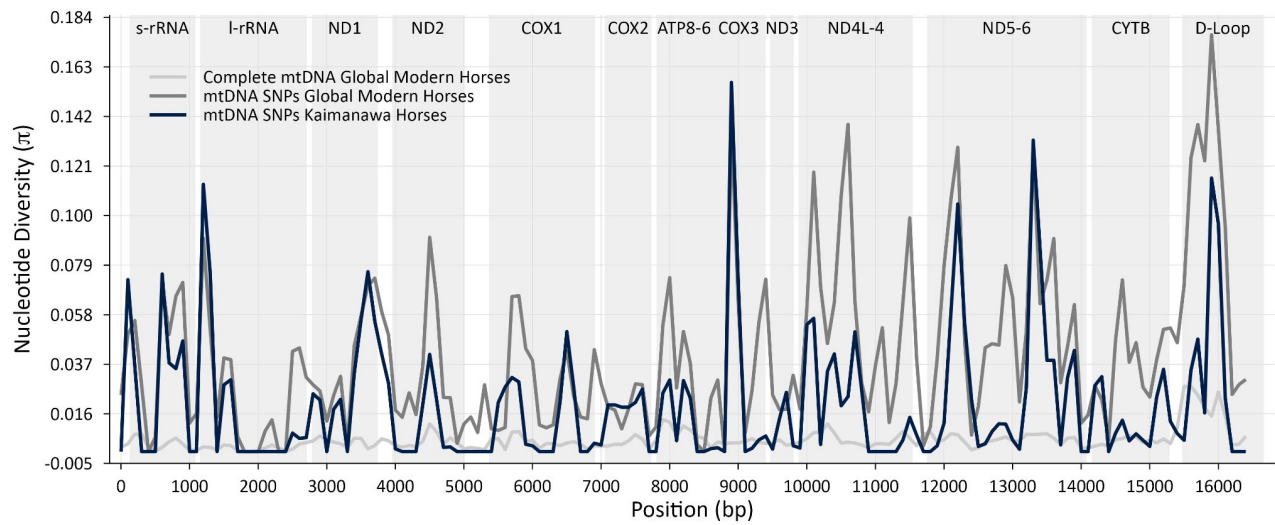


Figure S17. Window-based mitochondrial nucleotide diversity (π) across the maternally unrelated Kaimanawa Horses (blue, n SNPs = 1,078, n individuals = 101), global modern horses using SNPs (dark grey, n SNPs = 1,078, n individuals = 81), and global modern horses using the complete mtDNA dataset (light grey, n SNPs = 16,600, n individuals = 81). Mitochondrial regions are indicated in gray bars.