## **Supplementary information**

Genome-wide association study for frozen-thawed sperm motility in stallions across various horse breeds

Elena V. Nikitkina<sup>1</sup>, Natalia V. Dementieva<sup>1</sup>, Yuri S. Shcherbakov<sup>1</sup>, Mikhail M. Atroshchenko<sup>2</sup>, Andrei A. Kudinov<sup>1</sup>, Oleg I. Samoylov<sup>1</sup>, Marina V. Pozovnikova<sup>1</sup>, Artem P. Dysin<sup>1</sup>, Anna A. Krutikova<sup>1</sup>, Artem A. Musidray<sup>1</sup>, Olga V. Mitrofanova<sup>1</sup>, Kirill V. Plemyashov<sup>1</sup>, Darren K. Griffin<sup>3</sup>, and Michael N. Romanov<sup>3,4\*</sup>

Supplementary figure 1

## **Supplementary Figure**

Supplementary Figure S1. Distribution of the studied 23 horse breeds as built in the web tools Phantasus ([32]; A) and ClustVis ([33]; B) using sperm motility data before and after freezing. (A) PCA plot. Unit variance scaling is applied to rows; SVD with imputation is used to calculate principal components (PC). *X* and *Y* axes show PC1 and PC2 that explain 71.8% and 28.2% of the total variance, respectively. N = 23 data points (breeds). (B) Heatmap and clustering trees, with rows being centered and unit variance scaling being applied to rows. Both three rows and 23 columns are clustered using Euclidean distance and average linkage, with tightest cluster first for tree ordering.

