Does inbreeding contribute to pregnancy loss in Thoroughbred horses?

Title: Does inbreeding contribute to pregnancy loss in Thoroughbred horses?

Key words: Mare, Abortion, Miscarriage, Fetus, Homozygosity

Word count: 4,913

Authors: Jessica M. Lawson[1†*], Charlotte A. Shilton[2†], Victoria Lindsay-McGee[3‡], Androniki Psfidi[3], D. Claire Wathes[1], Terje Raudsepp[4], Amanda M. de Mestre[2§¶]

Author affiliations: 1) Department of Pathobiology and Population Sciences, The Royal Veterinary College, University of London, Hatfield, UK, 2) Department of Comparative Biomedical Sciences, The Royal Veterinary College, University of London, London, UK. 3) Department of Clinical Science and Services, The Royal Veterinary College, University of London, London, UK. 4) Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA.

† Jessica M. Lawson and Charlotte A. Shilton should be considered joint first author

‡ Current affiliation: The Royal (Dick) School of Veterinary Studies, The University of Edinburgh, UK

§Email of corresponding authors: jmlawson@rvc.ac.uk and amm43@cornell.edu

¶ Current affiliation: Baker Institute for Animal Health, Department of Biomedical Sciences, Cornell University, Ithaca, New York, USA

Declarations:

Authorship: A.M.dM., T.R. and D.C.W., were involved in the conceptualisation of the overall PhD project and acquisition of funding. TR provided the PCR protocol to sex the conceptuses. C.A.S. designed the specific experiment and C.A.S. and J.M.L. performed the formal analysis. C.A.S, J.M.L., and A.M.dM. prepared the original manuscript draft. V.L. and A.P. provided expert knowledge for the interpretation of genetic results generated. A.M.dM., D.C.W., T.R., V.L., and A.P. provided comments and edits of the original manuscript draft. T.R. oversaw the

supervision of the analysis of the SNP array and WGS methodologies, while A.M.dM. oversaw the supervision of other methodologies and the wider research activity. A.M.dM., C.A.S. and J.M.L take responsibility for the integrity of the data and the accuracy of the data analysis.

Source of Funding: Project grant awarded by the Thoroughbred Breeders Association, Horserace Betting Levy Board, the Alborada Trust, and partial PhD studentship funding from the Royal Veterinary College's Paul Mellon Trust for Equine Research.

Competing Interests: No competing interests.

Ethical Animal Research: All conceptus recoveries from clinical cases of pregnancy loss were performed with owner consent under ethics approval from the Clinical Research and Ethical Review Board at the Royal Veterinary College (URN:2012-1169 and URN:2017-1660-3). Animal care of the research herd from which the peripheral blood was obtained was in accordance with the Animals (Scientific Procedures) Act 1986 guidelines set by the Home Office and Ethics Committee of the Royal Veterinary College, London (HO licence PPL 70/8577).

Informed consent: Informed consent was obtained from all participating stud farms by means of information sheets regarding the use of samples and a signed consent form. Anonymity was maintained by coding the names of veterinarians, stud farms, mares, and stallions, with the codes maintained in a password protected Microsoft Excel database.

Acknowledgements: Anne Kahler , Belinda Rose and Daniel Hampshire for collecting and processing the EPLs.

Data availability statement: The genotyping data that support the findings of this study are not publicly available due to privacy and ethical consent restrictions, however, request for private access can be made to the corresponding authors.

Abstract

Background: Excessive inbreeding increases the probability of uncovering homozygous recessive genotypes and has been associated with an increased risk of retained placenta and lower semen quality. No genomic analysis has investigated the association between inbreeding levels and pregnancy loss. **Objectives**: This study compared genetic inbreeding coefficients (F) of naturally occurring Thoroughbred Early Pregnancy Loss (EPLs), Mid and Late term Pregnancy Loss (MLPL), and Controls. The F value was hypothesised to be higher in cases of pregnancy loss (EPLs and MLPLs) than Controls. **Study design**: Observational case-control study. **Methods**: Allantochorion and fetal DNA from EPL (n=37, gestation age 14-65 days), MLPL (n=94, gestational age 70 days–24 hours post parturition) and Controls (n=58) were genotyped on the Axiom Equine 670K SNP Genotyping Array. Inbreeding coefficients using Runs Of Homozygosity (FROH) were calculated using PLINK software. ROHs were split into size categories to investigate the recency of inbreeding. **Results**: MLPLs had significantly higher median number of ROH (188 interquartile range (IQR), 180.8-197.3), length of ROH (3.10, IQR 2.93-3.33), and total number of ROH (590.8, IQR 537.3-632.3), and F_{ROH} (0.26, IQR 0.24-0.28) when compared with the Controls and the EPLs ($p<0.05$). There was no significant difference in any of the inbreeding indices between the EPLs and Controls. The MLPLs had a significantly higher proportion of long (>10 Mb) ROH (2.5%, IQR 1.6-3.6) than the Controls (1.7%, IQR 0.6-2.5), p=0.001. No unique ROHs were found in the EPL or MLPL populations. **Limitations**: SNP-array data does not allow analysis of every base in the sequence. **Conclusions**: This first study of the effect of genomic inbreeding levels on pregnancy loss showed that inbreeding is a contributor to MLPL, but not EPL in the UK Thoroughbred population. Mating choices remain critical, because inbreeding may predispose to MLPL by increasing the risk of homozygosity for specific lethal allele(s).

1. Introduction

Inbreeding (the mating of related individuals) is a common practice in the livestock industry because individuals with desirable traits are highly prized as breeding stock. The descendants of these individuals therefore make up a greater proportion of the population. In the Thoroughbred breeding industry, with a focus on racing potential, 97% of 10,118 individuals studied could be traced to a single stallion, Norther Dancer [1]. The inbreeding coefficient (F) is the probability that a pair of alleles at a specific locus will be identical-by-descent [2] thus increasing the risk of uncovering undesirable recessive phenotypes. Historically, pedigree data have been used to estimate inbreeding levels, this however is limited particularly with by missing or incorrect data [3-5]. Relevant genomic estimations such as Runs Of Homozygosity (ROH) are considered preferable, and have the additional benefit of indicating inbreeding trends [5]. Over time, mutations break up longer ROH into shorter ROH, and thus short ROHs can estimate inbreeding that happened in the distant past, while longer ROH indicate a more recent occurrence of inbreeding.

In horses, inbreeding has been associated with an increased risk of retained placenta [6] and lower semen quality [7-10]. Fertility scores and foaling rates have been shown to have either no association [11; 12], a weak association [13], or a significant association [14] with inbreeding levels. Gestation length is not associated with inbreeding in horses [15-21]. Only a single study using pedigree data to calculate the inbreeding coefficient (FPED) has investigated any link between inbreeding and pregnancy loss, finding both increased FPED and mare age to be significant contributors to increased risk of early abortion at <5 months gestation in Norwegian Trotters [22]. To date no genomic analysis has been completed to determine any association between inbreeding levels and pregnancy loss in horses. Around 5-10% of equine pregnancies end in early pregnancy loss (EPL; up to 65 days gestation) [23], and a further 7.3% of equine pregnancies are lost between day 70 of gestation and 24 hours post parturition (mid

and late term pregnancy loss (MLPL)) [24]. The underlying causes of pregnancy loss differ between early and mid to late gestation [24-27]. A biobank of naturally occurring EPLs, created using recent advances in methodologies to collect tissue samples [28], and MLPLs, have allowed investigation of the FROH for the pregnancy itself rather than that of the parents. It was hypothesised that the inbreeding coefficient would be higher in cases of pregnancy loss (both EPLs and MLPL) than Controls. This project specifically aimed to compare the estimated genetic inbreeding coefficient using ROH between cases (naturally occurring EPLs and MLPL) and Controls.

2. Materials and Methods

2.1 Ethics statement

All conceptus recoveries from clinical cases of EPL and collections of MLPLs were performed with owner consent under ethics approval from the Clinical Research and Ethical Review Board of the research institute. Animal care of the research herd from which the peripheral blood was obtained was in accordance with the Animals (Scientific Procedures) Act 1986 guidelines set by the Home Office and Ethics Committee of the research institute. Informed consent was obtained from all participating stud farms by means of information sheets regarding the use of samples and a signed consent form. Anonymity was maintained by coding the names of veterinarians, stud farms, mares, and stallions, with the codes maintained in a password protected Microsoft Excel database.

2.2 Sample collection

Sample collection and processing of the EPLs has been reported previously [28; 29]. In brief, following confirmation of pregnancy failure before 65 days post ovulation (no heartbeat/collapsed vesicle), conceptuses were recovered by uterine lavage by the attending veterinarian during the 2013-2021 breeding seasons [28]. Successfully flushed conceptuses

were then placed in sterile transport media and stored at 4°C until being transported on ice to the laboratory for assessment and dissection within an hour of arrival. Placentae from cases of abortion, stillbirth, or perinatal death within 24 hours of parturition were obtained following submission for diagnostic investigation at a Newmarket based diagnostic laboratory during the 2017-2020 breeding seasons. Approximately 5x5 mm sections of allantochorionic tissue were taken and stored in 1.5 ml of DNAgard (Biometrica, Nevada, USA) and stored at room temperature for up to 6 months. When allantochorion was not available, sections of fetal gluteal muscle measuring 5x5 mm were dissected and stored following the same protocol.

The control group were adult UK Thoroughbreds (n=58 mares, all over 3 years old). Peripheral Blood Mononuclear Cells (PBMCs) were isolated from whole blood following collection from the jugular vein of Thoroughbred mares (n=5) from the institutional research herd as previously described [30]. PBMC pellets were then snap frozen in liquid nitrogen and transferred to -80°C. Hair samples from 53 Thoroughbreds, across eight UK stud farms, were submitted anonymously by the attending veterinarians between 2017 and 2021. The eight stud farms represented a sub population of the stud farms which had submitted EPLs and MLPLs. Aside from the name of the stud farm that the sample came from, no additional clinical data was collected beyond the individual fitting the criteria of being a registered Thoroughbred, over three years of age.

2.3 DNA extraction

DNA from frozen tissues, tissue stored in DNAgard and PBMCs were extracted using QIAGEN DNeasy Blood and Tissue kit (Qiagen Sciences, Maryland, USA), following manufacturer's guidelines. Briefly, tissue or cells were incubated at 56°C overnight in 180 µl buffer ATL and 20 µl proteinase K (600 mAU/ml). Tissues were then incubated at room temperature for 2 minutes with 28 U RNase A as recommended by the manufacturer then passed through a spin column, before elution with 100 µl Buffer AE provided in the kit.

Intact roots from 15 hairs were lysed in a mix of 300 µl cell lysis solution and 5 µl proteinase K at 37 °C overnight. To the supernatant, 100 µl protein precipitation solution (PPS) was added then vortexed and incubated on ice for 10 minutes. Following centrifugation for 3.5 minutes at 16,000g, the supernatant was added to 300 µl isopropanol and mixed by inverting 40 times, then centrifuged again for 3.5 minutes at 16,000g. The supernatant was discarded and 300 µl 70% ethanol added, vortexed for 45 seconds, then centrifuged for 3.5 minutes at 16,000g. The supernatant was discarded, and the pellet dried overnight at room temperature. Once dried, the pellet was resuspended in 20 µl of hydration solution.

All DNA was quantified using a DeNovix Spectrophotometer (DeNovix, Delaware, USA), measuring quantity (ng/µl) and quality (A260/A230 and A260/A280). DNA quality was confirmed to have no effect on the inbreeding values calculated.

2.4 Genotype preparation and SNP pruning

The resulting .CEL files generated from all samples (Cases and Controls) hybridised to the Axiom™ Equine 670K SNP Genotyping Array were imported into Axiom Analysis Suite (AxAS, v5.0.1.38), with SNP probe locations based on EquCab3.0 reference genome. Following the "Genotyping" workflow, genotype data were exported as a .vcf file. SNP quality control (QC) settings were kept as default as recommended by the manufacturer. Only SNP probes that met AxAS "Best and Recommended" (i.e. passed all internal programme QC metrics) were included in the exported .vcf file.

As there appears to be little consensus on the filtering steps required for Runs Of Homozygosity (ROH) analyses, SNPs were not filtered based on Hardy-Weinberg Equilibrium (HWE), Minor Allele Frequency (MAF), or Linkage Disequilibrium (LD), the latter two being in accordance with recently published guidelines [31]. The removal of rare variants may artificially inflate or deflate calls, potentially missing critical ROHs. To reduce the calling of

ROHs that were in LD, the minimum length of ROH was set to 1 Mb for analysis of groups within this study. Only diploid samples were tested, with aneuploid and euploid individuals removed prior to analysis.

2.5 Runs of Homozygosity (ROH) detection in PLINK

The .vcf files generated above were then used to identify ROHs in PLINK v1.90 [32] using the options as previously described [33; 34]. The options used were as follows: minimum SNP density = one SNP per 50 kb, maximum gap length = 100 kb, minimum length per ROH = 1 Mb, minimum number of homozygous SNPs = 80, maximum number of heterozygous SNPs per ROH = 1, maximum number of missing SNPs per ROH = 2. Only autosomes were included in this analysis.

2.6 ROH analysis

The total number of ROHs per individual (N_{ROH}), the average length ROH an individual possessed (L_{ROH}), and the total length of all ROHs (S_{ROH}) were next calculated for each sample using the outputs generated in PLINK. The genomic inbreeding coefficient (F_{ROH}) was calculated by dividing the S_{ROH} by the total autosomal genome length (L_{AUTO}; [35]). The autosomal length for EquCab3.0 was 2,281,300 kb (2,280.9 Mb) as calculated from values on ENSEMBL (http://www.ensembl.org/Equus_caballus/Location/Chromosome?r=25%3A1-1000).

$$FROH = \frac{SROH}{LAUTO}$$

The number of short- (1-2 Mb) and long (>10 Mb) ROHs (in similarity with [33]) for each category were calculated per individual and the percentage of ROHs in each group per individual were then compared between groups.

2.7 Unique ROH

To identify any candidate ROHs associated with pregnancy loss, ROHs detected in EPLs, and separately MLPLs, were combined into .csv files and compared to all ROHs detected in Controls using *bedtools intersect* pipeline.

2.8 Statistical analysis

Normality of the data was assessed in GraphPad Prism (v9.1.2, https://www.graphpad.com/) using the Shapiro-Wilk normality test. In all cases, the normality tests failed, and therefore Kruskal-Wallis with post hoc Dunn's test were used to identify statistical differences between groups, with significance set at $p<0.05$. The median and interquartile ranges are presented throughout.

3. Results

3.1 ROH differ between mid and late pregnancy losses and Controls

The EPLs (n=37) were obtained from 22 stud farms and the observed gestation ages ranged from 14-68 days. The MLPLs (n=94) came from 42 stud farms, for 9/94 cases the stud farm was unavailable. The observed gestational age in the MLPL group ranged from 86 days gestation to 24 hours post parturition. The ROHs of EPLs and MLPLs were compared to each other and with Controls (n=58). The EPLs and Controls did not significantly differ in median values of N_{ROH}/L_{ROH}/ S_{ROH}/F_{ROH} (Figure 1.a-d).

The MLPLs had significantly higher values for all four metrics (N_{ROH} = 188 (IQR 180.8-197.3), L_{ROH} = 3.10 (IQR 2.93-3.33), S_{ROH} = 590.8 (IQR 537.3-632.3), and F_{ROH} = 0.26 (IQR = 0.24-0.28) when compared with both the EPLs and the Controls, $p<0.05$ (Figure 1.a-d). MLPLs were further explored as Abortions (70-300 days of gestation, n=74) and Stillbirths (301 days of gestation to 24 hours post parturition, n=16) and no significant difference found between the groups in any of the inbreeding indices, $p>0.05$. Four MLPLs were excluded from

this additional analysis as, although they could be categorised as a MLPL based on the crown rump length of the fetus, only estimated gestational ages were available.

3.2 MLPLs show a higher degree of recent inbreeding

Previous work has shown that shorter ROHs (smaller than 0.5 Mb) are indicative of historical inbreeding from 50-100 generations ago i.e., before the establishment of the Thoroughbred breed [36]. The EPLs had a significantly lower median percentage of short ROH (1-2 Mb; 47.3%, IQR 42.1-50.2) than the Controls (48.1%, IQR 46.5-54.3), p=0.02, Figure 2.a). The MLPL had significantly higher percentages of long ROH (>10 Mb; 2.5%, IQR 1.6-3.6) compared to the Controls (1.7%, IQR 0.6-2.5), p=0.001 (Figure 2.b), but were not significantly different from the EPLs (1.8%, IQR 1.2-3.1), p=0.3. There was no significant difference in the percentages of short length ROH between MLPLs (48.8, IQR 44.8-51.8) and either EPLs (47.3%, IQR 42.1-50.2) or Controls (48.1, IQR 46.5-54.3), p=0.2 and 0.8 respectively (Figure 2.a).

3.3 No ROHs were found to be specific to pregnancy loss

In total, 9,682 ROHs were found across 58 Controls, 6,460 ROHs were found across 37 EPLs, and 16,395 ROHs were found across 94 MLPL. To investigate whether specific ROHs may be lethal, the ROH call lists from EPLs, and separately the MLPLs, were compared with the ROH call list from the Controls. No ROH calls came up as unique between the EPL and Controls, the MLPL and Controls, or the EPLs and MLPLs.

4. Discussion

Approximately 5-15% of confirmed equine pregnancies fail before 65 days of gestation, with a further 7.3% failing before the end of the first day of life [24]. To date, no study has

specifically investigated any link between genetic inbreeding metrics and pregnancy loss in the mare. This study found that pregnancies lost in mid and late gestation (MLPLs), from Thoroughbred mares in the UK, had significantly higher inbreeding metrics than UK adult Thoroughbred horses, with the proportion of long ROH (an indicator of recent inbreeding) also increased in these lost pregnancies. Contrary to the initial hypothesis, pregnancies lost early in gestation (EPLs) were found to show no significant difference in inbreeding metrics compared to UK adult Thoroughbred horses. No ROHs were found to be unique to the EPL or MLPL cohort.

Higher inbreeding metrics will be associated with an increased risk of the individual inheriting a deleterious homozygous mutation. Examples of homozygous single point mutations that are known to result in pregnancy loss and other congenital abnormalities include congenital hepatic fibrosis [37], congenital hydrocephalus [38], and warmblood fragile foal syndrome [39], the latter recently described as a cause of pregnancy loss for the first time in a Thoroughbred [40]. The findings of our study further underpin the importance of continued research into identifying and characterising fatal mutations, and with new mutations arising all the time, continued surveillance is important. SNP mutations have been associated with abortion and stillbirth [38-40], but to date none have been identified as causes of lethality in EPLs. Whilst the presence of defective recessive alleles in homozygous status could still contribute to EPL as a less common or rare phenomena, our data support the hypothesis that SNP mutations are more likely to cause lethality in mid to late gestation. It should also be noted that as we only explored diploid cases, we cannot understand the effects of inbreeding on aneuploidy and other chromosomal abnormalities from this data.

While inbreeding theoretically increases the risk of the offspring inheriting the same deleterious mutation from both parents, practically the link may not be as linear as expected. Thoroughbreds were ranked 3rd amongst 37 horse breeds for inbreeding coefficient but 9th for

genomic mutational load (genetic burden due to accumulation of deleterious mutations) [41]. The protein-coding mutational load is even more nuanced, with almost all the 37 breed groups studied overlapping, regardless of their inbreeding levels. The relatively lower mutational load of Thoroughbreds may in part be due to the breeding practice of selecting for racing potential. Individuals born with a poor phenotype would either not enter racing or have a poor performance on the track so would be unlikely to enter the breeding stock. Likewise, MLPL may act as a successful natural purging step, preventing the individual from entering the national herd in the first place and reproducing.

The MLPLs were found to have a significantly higher proportion of long (>10 Mb) ROH than the Controls. Longer ROHs are indicative of more recent inbreeding as consanguineous matings are more likely to share a greater number of alleles. Over time, heterozygosity can be reintroduced to the population through mutations which break up ROHs into smaller runs. The presence of the higher percentage of long ROH in the MLPL group follows the same trend as the regression analysis of FROH over five decades by McGivney et al [1]. In Great Britain the number of stallions registered for covering has almost halved in the last 10 years, from 285 stallions in 2011 to 147 stallions in 2021.[42; 43], restricting the choice for breeders. Whilst Thoroughbred breeders make careful selection of their matings and breeding choices, the effects of this decline should be under continued scrutiny by the industry to prevent the Thoroughbred populations from suffering an inbreeding depression.

There is limited comparative data available. Todd et al. [20] explored inbreeding levels using pedigree data of Australian Thoroughbreds and found no significant association of the mare, stallion or conceptus' inbreeding coefficients with the foaling rate. Klemetsdal and Johnson [22] also used pedigree data, this time in Norwegian Trotters, and observed that the inbreeding coefficient of the potential offspring (i.e. the pregnancy) was not a significant contributor to foaling rate (proportion of covers resulting in a live foal) in their modelling.

298 Klemetsdal and Johnson [22] also explored predictors of early abortion (pregnancy loss prior
299 to month 5 of gestation). Whilst they reported that a 1% increase in a mare's inbreeding
300 coefficient was associated with a 1.27% increase in early abortion frequency, they found that
301 the inbreeding coefficient of the pregnancy itself was not significantly associated with early
302 abortion [22]. This study period only partially overlaps the phenotypes we explored and uses
303 pedigree derived inbreeding coefficients rather than genomic data. Our data suggests that
304 inbreeding exerts an effect on pregnancy loss from day 70 of gestation all the way through to
305 24 hours post parturition.

306 Thoroughbreds were ranked 3rd amongst 37 horse breeds for inbreeding coefficient but
307 9th for genomic mutational load (genetic burden due to accumulation of deleterious mutations)
308 [41]. The protein-coding mutational load is even more nuanced, with almost all the 37 breed
309 groups studied overlapping, regardless of their inbreeding levels. The relatively lower
310 mutational load of Thoroughbreds may in part be due to the breeding practice of selecting for
311 racing potential. Individuals born with a poor phenotype would either not enter racing or have
312 a poor performance on the track so would be unlikely to enter the breeding stock. In the same
313 way, MLPL may be a successful natural purging step, preventing the individual from entering
314 the national herd in the first place and reproducing.

315 There are limitations to this study as the sample sizes are relatively small, it is restricted
316 to one breed and we used a mixture of DNA sources, from the placenta and fetus in the
317 pregnancy losses, and hair and PBMC in the Controls. The samples were submitted from
318 numerous stud farms across the UK; however, it is acknowledged that self-selection bias in the
319 farms and veterinarians who chose to submit material may affect the results. Further, the cause
320 of the loss may in some cases, reduce or preclude the availability of tissue, for example
321 sampling of EPLs is reliant on products of conception being available for collection via uterine
322 lavage and submitted for analysis. Similarly, some causes of MLPL may not be submitted to a

diagnostic laboratory for post mortem examination if investigation is not perceived to be required, for example an intrapartum stillbirth from distal limb contractions. This opens up the potential for bias in the phenotypes assessed in this study. Further, non-diploid EPL and MLPL samples were excluded from the analysis due to the possibility of inflated or reduced F_{ROH} coefficients related to the ploidy status that could have impacted the results. Given the high proportion of chromosome wide copy number variants in EPLs [29], this would have disproportionately affected this phenotype and be a source of bias. It would be of interest to repeat this analysis with different breeds and with larger sample sizes. Whilst other factors such as year of sampling and DNA quality could plausibly impact the results, the inclusion of multiple breeding seasons, the exclusion of failed probes and only individuals that had a SNP call rate of > 98% will have minimised their influence.

In conclusion, we observed higher inbreeding metrics in UK Thoroughbred pregnancies lost in mid and late gestation compared to the adult population, evidencing that lack of heterogeneity is a contributor to pregnancy failure after the early pregnancy period. We hypothesize that this is due to an increase in the occurrence of homozygous recessive alleles, highlighting that studies into the role of specific gene mutations are both required and warranted. Although no significant differences were observed in the inbreeding metrics between the EPL and the UK Thoroughbred adults, we recognise a bias in the phenotypes of the losses in this group. Our data highlights the importance of cognisance in mating decisions in the Thoroughbred industry, and continued work in the laboratory to identify possible deleterious mutations.

2. Manufacturer's addresses

Axiom™ Equine Genotyping Array (Axiom MNEC670) - ThermoFisher Scientific

Axiom Analysis Suite (AxAS, v5.0.1.38) - ThermoFisher Scientific

GraphPad Prism (v9.1.2) - GraphPad

3. Tables

No tables

4. Figure Legends

, **Figure 1: Runs of homozygosity (ROH) analysis of Thoroughbred EPLs and MLPLs compared with breed matched Controls.** Thoroughbred Early Pregnancy Losses (EPL, n=37) were compared with Mid and Late Term Pregnancy Losses (MLPL, n=94) and breed matched Controls (n=58). The MLPLs had significantly increased median a) NROH = number of ROH ($p<0.001$), b) SROH = sum of ROH ($p<0.001$), c) LROH = average length of ROH ($p<0.001$), and d) FROH = inbreeding coefficient ($p<0.001$). Black line = median and interquartile range. Kruskal-Wallis with Dunn's multiple comparisons test (* =$p<0.05$, ** = $p<0.01$, *** = $p<0.001$).

Figure 2: Runs of homozygosity (ROH) size analysis of Thoroughbred EPLs, MLPLs and breed matched Controls. Early Pregnancy Loss (EPL; n=37), Mid and Late Term Pregnancy Loss (MLPL, n=94) and breed matched Controls (n=58) and the proportion of a) short- (1-2 Mb), and b) long- (>10 Mb) ROHs. Black line = median and interquartile range. Kruskal-Wallis with Dunn's multiple comparisons test (* $p=0.02$, ** $p=0.001$).

5. List of legends for Supplementary items

No supplementary items

6. References

[1] McGivney, B.A., Han, H., Corduff, L.R., Katz, L.M., Tozaki, T., MacHugh, D.E. and Hill, E.W. (2020) Genomic inbreeding trends, influential sire lines and selection in the global Thoroughbred horse population. *Scientific reports* **10**, 1-12.

[2] Gurgul, A., Szmatoła, T., Topolski, P., Jasielczuk, I., Żukowski, K. and Bugno-Poniewierska, M. (2016) The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle. *Journal of applied genetics* **57**, 527-530.

[3] Alemu, S.W., Kadri, N.K., Harland, C., Faux, P., Charlier, C., Caballero, A. and Druet, T. (2021) An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity* **126**, 410-423.

[4] Howrigan, D.P., Simonson, M.A. and Keller, M.C. (2011) Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC genomics* **12**, 1-15.

[5] Keller, M.C., Visscher, P.M. and Goddard, M.E. (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* **189**, 237-249.

[6] Sevinga, M., Vrijenhoek, T., Hesselink, J., Barkema, H. and Groen, A. (2004) Effect of inbreeding on the incidence of retained placenta in Friesian horses. *J Anim Sci* **82**, 982-986.

[7] Aurich, C., Achmann, R. and Aurich, J.E. (2003) Semen parameters and level of microsatellite heterozygosity in Noriker draught horse stallions. *Theriogenology* **60**, 371-378.

[8] Van Eldik, P., Van Der Waaij, E., Ducro, B., Kooper, A., Stout, T. and Colenbrander, B. (2006) Possible negative effects of inbreeding on semen quality in Shetland pony stallions. *Theriogenology* **65**, 1159-1170.

[9] Dini, P., Bartels, T., Revah, I., Claes, A.N., Stout, T.A. and Daels, P. (2020) A retrospective study on semen quality parameters from four different Dutch horse breeds with different levels of inbreeding. *Theriogenology* **157**, 18-23.

[10] Ducro, B. (2011) *Relevance of test information in horse breeding*, Wageningen University and Research.

[11] Mahon, G. and Cunningham, E. (1982) Inbreeding and the inheritance of fertility in the thoroughbred mare. *Livestock Production Science* **9**, 743-754.

[12] Müller-Unterberg, M., Wallmann, S. and Distl, O. (2017) Effects of inbreeding and other systematic effects on fertility of Black Forest Draught horses in Germany. *Acta Veterinaria Scandinavica* **59**, 1-6.

[13] Cothran, E., MacCluer, J., Weitkamp, L., Pfennig, D. and Boyce, A. (1984) Inbreeding and reproductive performance in Standardbred horses. *Journal of heredity* **75**, 220-224.

[14] Sairanen, J., Nivola, K., Katila, T., Virtala, A.-M. and Ojala, M. (2009) Effects of inbreeding and other genetic components on equine fertility. *Animal* **3**, 1662-1672.

[15] Valera, M., Blesa, F., Dos Santos, R. and Molina, A. (2006) Genetic study of gestation length in Andalusian and Arabian mares. *Anim Reprod Sci* **95**, 75-96.

[16] Gonçalves, R.W., da Costa, M., Rocha, J., da Costa, M., Silva, E. and Ribeiro, A. (2011) Inbreeding effect on reproductive traits in a herd of Mangalarga Marchador Brazilian horses. *Revista Brasileira de Saúde e Produção Animal* **12**, 641-649.

[17] Bene, S., Benedek, Z., Szabo, F. and Polgár, P. (2014) Some effects on gestation length of traditional horse breeds in Hungary. *Journal of Central European Agriculture* **15**, 0-0.

[18] Christmann, A., Sieme, H., Martinsson, G. and Distl, O. (2017) Genetic and environmental factors influencing gestation length and parturition conception interval in Hanoverian Warmblood. *Livestock Science* **199**, 63-68.

[19] Ewert, M., Lüders, I., Böröcz, J., Uphaus, H., Distl, O. and Sieme, H. (2018) Determinants of gestation length in Thoroughbred mares on German stud farms. *Anim Reprod Sci* **191**, 22-33.

[20] Todd, E.T., Hamilton, N.A., Velie, B.D. and Thomson, P.C. (2020) The effects of inbreeding on covering success, gestation length and foal sex ratio in Australian thoroughbred horses. *BMC genetics* **21**, 1-9.

[21] Rodrigues, J.A., Gonçalves, A.R., Antunes, L., Bettencourt, E.V. and Gama, L.T. (2020) Genetic and environmental factors influencing gestation length in Lusitano horses. *J Equine Vet Sci* **84**, 102850.

[22] Klemetsdal, G. and Johnson, M. (1989) Effect of inbreeding on fertility in Norwegian trotter. *Livestock production science* **21**, 263-272.

[23] Rose, B., Firth, M., Morris, B., Roach, J., Wathes, D., Verheyen, K. and De Mestre, A. (2018) Descriptive study of current therapeutic practices, clinical reproductive findings and incidence of pregnancy loss in intensively managed thoroughbred mares. *Anim Reprod Sci* **188**, 74-84.

[24] Roach, J.M., Foote, A.K., Smith, K.C., Verheyen, K.L. and Mestre, A.M. (2021) Incidence and causes of pregnancy loss after Day 70 of gestation in Thoroughbreds. *Equine veterinary journal* **53**, 996-1003.

[25] Shilton, C.A., Kahler, A., Roach, J.M., Raudsepp, T. and de Mestre, A.M. (2023) Lethal variants of equine pregnancy: is it the placenta or foetus leading the conceptus in the wrong direction? *Reproduction, Fertility and Development* **35**, 51-69.

[26] Hamstead, L., Chang, Y.-M., Crowhurst, J., Wise, Z., McGladdary, A., Ricketts, S. and de Mestre, A.M. (2012) Retrospective study of early pregnancy loss in Thoroughbred mares. *Equine veterinary journal* **44**, 2-18.

[27] Macleay, C.M., Carrick, J., Shearer, P., Begg, A., Stewart, M., Heller, J., Chicken, C. and Brookes, V.J. (2022) A Scoping Review of the Global Distribution of Causes and Syndromes Associated with Mid-to Late-Term Pregnancy Loss in Horses between 1960 and 2020. *Veterinary Sciences* **9**, 186.

[28] Rose, B.V., Cabrera-Sharp, V., Firth, M.J., Barrelet, F.E., Bate, S., Cameron, I.J., Crabtree, J.R., Crowhurst, J., McGladdery, A.J., Neal, H., Pynn, J., Pynn, O.D., Smith, C., Wise, Z.,

Verheyen, K.L., Wathes, D.C. and de Mestre, A.M. (2016) A method for isolating and culturing placental cells from failed early equine pregnancies. *Placenta* **38**, 107-111.

[29] Shilton, C.A., Kahler, A., Davis, B.W., Crabtree, J.R., Crowhurst, J., McGladdery, A.J., Wathes, D.C., Raudsepp, T. and De Mestre, A.M. (2020) Whole genome analysis reveals aneuploidies in early pregnancy loss in the horse. *Scientific Reports* **10**, 13314.

[30] Robbin, M.G., Wagner, B., Noronha, L.E., Antczak, D.F. and de Mestre, A.M. (2011) Subpopulations of equine blood lymphocytes expressing regulatory T cell markers. *Veterinary immunology and immunopathology* **140**, 90-101.

[31] Meyermans, R., Gorssen, W., Buys, N. and Janssens, S. (2020) How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC genomics* **21**, 1-14.

[32] Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I. and Daly, M.J. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics* **81**, 559-575.

[33] Grilz-Seger, G., Mesarič, M., Cotman, M., Neuditschko, M., Druml, T. and Brem, G. (2018) Runs of homozygosity and population history of three horse breeds with small population size. *J Equine Vet Sci* **71**, 27-34.

[34] Druml, T., Neuditschko, M., Grilz-Seger, G., Horna, M., Ricard, A., Mesarič, M., Cotman, M., Pausch, H. and Brem, G. (2018) Population networks associated with runs of homozygosity reveal new insights into the breeding history of the Haflinger horse. *Journal of heredity* **109**, 384-392.

[35] McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O. and Tenesa, A. (2008) Runs of homozygosity in European populations. *The American Journal of Human Genetics* **83**, 359-372.

[36] Fawcett, J.A., Sato, F., Sakamoto, T., Iwasaki, W.M., Tozaki, T. and Innan, H. (2019) Genome-wide SNP analysis of Japanese Thoroughbred racehorses. *PLoS One* **14**, e0218407.

[37] Drögemüller, M., Jagannathan, V., Welle, M.M., Graubner, C., Straub, R., Gerber, V., Burger, D., Signer-Hasler, H., Poncet, P.-A. and Klopfenstein, S. (2014) Congenital hepatic fibrosis in the Franches-Montagnes horse is associated with the polycystic kidney and hepatic disease 1 (PKHD1) gene. *PloS one* **9**, e110125.

[38] Ducro, B.J., Schurink, A., Bastiaansen, J.W., Boegheim, I.J., van Steenbeek, F.G., Vos-Loohuis, M., Nijman, I.J., Monroe, G.R., Hellinga, I. and Dibbits, B.W. (2015) A nonsense mutation in B3GALNT2 is concordant with hydrocephalus in Friesian horses. *BMC genomics* **16**, 1-9.

[39] Aurich, C., Müller-Herbst, S., Reineking, W., Müller, E., Wohlsein, P., Gunreben, B. and Aurich, J. (2019) Characterization of abortion, stillbirth and non-viable foals homozygous for the Warmblood Fragile Foal Syndrome. *Anim Reprod Sci* **211**, 106202.

[40] Grillos, A., Roach, J., de Mestre, A., Foote, A., Kinglsey, N., Mienaltowski, M. and Bellone, R. (2021) First reported case of fragile foal syndrome type 1 in the Thoroughbred caused by PLOD1 c. 2032G> A. *Equine veterinary journal*.

[41] Orlando, L. and Librado, P. (2019) Origin and evolution of deleterious mutations in horses. *Genes* **10**, 649.

[42] Weatherbys (2020) Weatherbys Factbook 2020, Weatherbys GSB Ltd.

[43] Authority, B.H. (2014) *Fact Book 2014*, British Horseracing Authority, London, UK. p 49.