

Short Communication

Relatedness and genomic inbreeding in a sample of Timor ponies

Doris E. Fröhlich^a, Barbara Wallner^b, Rytis Juras^c, E. Gus Cothran^c, Brandon D. Velie^{a,*}^a Equine Genetics and Genomics Group, School of Life and Environmental Sciences, University of Sydney, NSW, Australia^b Institute of Animal Breeding and Genetics, Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Vienna, Austria^c Department of Veterinary Integrative Biosciences, School of Veterinary Medicine and Biomedical Science, Texas A&M University, College Station, TX 77843, United States

ARTICLE INFO

Keywords:

Breeding program
Genomics
Horse management
Inbreeding
Relationship

ABSTRACT

Timor ponies (TP) were first shipped to Australia in the early 1800s and were highly valued as transport and pack animals, which resulted in TPs contributing to the development of Australian horse breeds. Today, while the exact number of TPs in Australia is currently unknown, there has been recent interest in establishing a domestic breeding program for Australian TPs. The aim of this study was to evaluate the relatedness of a sample of TPs, as well as provide estimates of genomic inbreeding levels to better inform the feasibility of using these animals as founders for a domestic breeding program. Hair samples from each horse were genotyped using the Illumina 80K Infinium Equine genotyping array and data were analysed using PLINK v1.90b7, KING 2.3.2 and R v4.3.1. The results illustrate that there are distantly related and minimally inbred horses within the sampled TPs. Lengths of the ROH segments also indicated that recent inbreeding events are likely to only have occurred in a third of the horses. Overall, these results are promising for the success of a domestic TP breeding program; however, considering the low number of domestic TPs known to reside in Australia, there would certainly still be substantial benefits to incorporating additional TPs either directly from Timor or from areas in Australia that are believed to contain wild descendants of TPs.

1. Introduction

Timor ponies (TP) are a type of horse originally developed in Timor, an island north of Australia [1,2]. In the early 1800s, TPs were shipped to Australia as these horses were readily available to the early northern settlements and were better suited to the harsh Australian climate compared to traditional European horse breeds (e.g. Thoroughbreds, Warmbloods) [1,3]. Although TPs are only 100 to 120 cm high at withers, they are described as strong, resilient, willing, and speedy [1,4]. Consequently, this type of horse was historically used for transport, packing, and pony races [5] and it is this multi-purpose ability that resulted in TPs playing a role in the development of Australian horse stock, such as the Waler Pony and the Brumby [1]. While the exact number of TPs in Australia is currently unknown, there has been recent interest in establishing a domestic breeding program for Australian TPs. As such, the aim of the current study was to evaluate the relatedness of a small sample of TPs, as well as provide estimates of genomic inbreeding levels to better inform the feasibility of using these horses (H) as founders for a domestic breeding program.

2. Material and methods

Hair samples were collected from two captured TP stallions (H2, H11) and one captured TP mare (H3) from the Cobourg Peninsula, as well as five suspected descendants of these horses (H4-H6, H8-H9) and 15 TPs with varying origins. These 15 TPs were either from Timor and the surrounding islands (H7, H12-H15, H17-H23) or from unknown origins (H10, H16, H24) (Table 1). Additionally, a sample from a known descendant (H1), which is the offspring of H2 and H3, was included. Furthermore, H2, H3, H7 and H10-H24 were previously confirmed as TP by breed testing performed by the current owners of these horses and/or as part of a previous study [6]. All animal procedures were approved by the University of Sydney's Animal Ethics Committee (Project Number 2022/2138). Deoxyribonucleic acid (DNA) was extracted from all samples using standard protocols for DNA extraction from hair samples [6]. DNA from each horse was then genotyped using the Illumina 80K Infinium Equine genotyping array. Genotype data were analysed using PLINK v1.90b7, KING 2.3.2, and R v4.3.1 [7,8,9]. Quality control (QC) was performed to remove poorly genotyped and noisy data wherein

* Corresponding author at: Equine Genetics & Genomics Group, School of Life and Environmental Sciences, University of Sydney, RMC Gunn B19-603, NSW 2006, Australia.

E-mail address: brandon.velie@sydney.edu.au (B.D. Velie).

<https://doi.org/10.1016/j.jevs.2024.105016>

Received 28 November 2023; Received in revised form 22 January 2024; Accepted 26 January 2024

Available online 27 January 2024

0737-0806/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Table 1
Origins of the Timor ponies analysed in the current study.

Horse	Origin
H1	Captive bred and born
H2	Cobourg Peninsula, NT, AUS
H3	Cobourg Peninsula, NT, AUS
H4	Queensland, AUS
H5	Queensland, AUS
H6	Queensland, AUS
H7	Timor and surrounding islands
H8	Northern Territory, AUS
H9	Queensland, AUS
H10	Unknown
H11	Cobourg Peninsula, NT, AUS
H12	Timor and surrounding islands
H13	Timor and surrounding islands
H14	Timor and surrounding islands
H15	Timor and surrounding islands
H16	Unknown
H17	Timor and surrounding islands
H18	Timor and surrounding islands
H19	Timor and surrounding islands
H20	Timor and surrounding islands
H21	Timor and surrounding islands
H22	Timor and surrounding islands
H23	Timor and surrounding islands
H24	Unknown

single nucleotide polymorphisms (SNPs) with a call rate lower than 0.9, and SNPs located on chromosomes X and Y were excluded. For estimation of relationship coefficients (r), SNPs with minor allele frequency (MAF) < 0.05 were also removed [10,11]. Relationship coefficients, as the proportion of the genome shared between two individuals based on identical by descent (IBD), were estimated in both PLINK and KING [7,8,12]. The proportion of IBD for each horse pair was evaluated using the “genome” command in PLINK and the “ibdseg” command in KING. Two programs were used to estimate relationship coefficients (r) as each program utilizes a different estimation approach [7,8]. The calculation in PLINK is based on the allele frequencies and is computed by the equation:

$$r = \frac{1}{2}P(\text{IBD}=1) + P(\text{IBD}=2) \quad (1)$$

where $P(\text{IBD}=1)$ and $P(\text{IBD}=2)$ are the probability of sharing 1 or 2 alleles IBD for two individuals from the same random-mating population [7]. KING, on the other hand, utilizes the number of loci at which the genotypes of two individuals are heterozygotes ($N_{1,1}$) or different homozygotes ($N_{2,0}$), with r estimated as:

$$r = 2 * \frac{N_{1,1} - N_{2,0}}{N_1^{(i)} + N_1^{(j)}} \quad (2)$$

where $N_1^{(i)}$ and $N_1^{(j)}$ are the sum of the heterozygous loci of each individual [8]. Both programs reportedly have high accuracy for detecting first- and second-degree relationships, but accuracy decreases for distant relationships, with PLINK tending to underestimate and KING to overestimate distant relatedness [13]. The classification recommended by Manichaikul et al. (2010) was used to derive the degree of relatedness from the estimated r [8]. ROH analyses were carried out in PLINK using the “homozyg” command [7]. Parameter settings were chosen according to Meyermans et al. (2020), therefore minimal density for segments was set to 38 kb/SNP, maximal gap length was 400 kb, minimal number of SNPs for segments was calculated by the L-parameter and set to 57, and scanning window threshold was 0.052 [14]. Additionally, parameter settings included: length for segments ≥ 500 kb, number of heterozygotes allowed in a window = 0, and number of missing calls allowed in a window = 1 [11,15,16]. Custom scripts in R were then used to determine ROH islands that were present in at least 50% of the analysed individuals [9,17]. ROH islands were compared to previously

characterized quantitative trait loci (QTL) for reproduction and health traits in the horse to identify any overlap with regions associated with health traits or reproduction traits [18].

3. Results

In total, 24 individuals ($n=9$ female, $n=15$ male) were included in the final analyses. 74,251 SNPs passed QC prior to filtering for MAF and 56,916 SNPs passed QC after the MAF threshold was included. Known parent-offspring relationships between H1 and H1's sire (H2) and dam (H3) were validated in both PLINK and KING (Fig. 1). There was also evidence to support clear first- and second-degree relationships between the five suspected descendants (H4-H6, H8-H9), one TP with unknown origin (H10), and all three of the original TPs from the Cobourg Peninsula (H2, H3, H11). Overall, analyses in PLINK resulted in 17 first-, 54 second-, 51 third-, and 9 fourth-degree relationships, whereas the analyses in KING detected 10 first-, 35 second-, 42 third-, and 75 fourth-degree relationships. The remaining horse pairs were distantly related (more than four degrees) or unrelated and made up 53% and 41% of all horse pairs, respectively. As expected, results from both PLINK and KING revealed two groups of horses (Group 1 = H1-H3, H6, H8-H11, Group 2 = H12-H22), where Group 1 (G1) included the three TPs from the Cobourg Peninsula, and Group 2 (G2) consisted of 11 out of 15 TPs from varying origins. Relationship estimates are consistent with the fact that horses in G1 are from the same “family”, while the origins of the horses in G2 are more scattered. Additionally, five horses were unable to be allocated to either G1 or G2. Two of these (H23, H24) show only minimal relatedness with all horses analysed. The other three horses (H4, H5, H7) show a high level of relatedness to horses in both G1 and G2, with H4 and H5 being more closely related to horses from G1 and H7 being more closely related to horses from G2.

For the calculation of inbreeding coefficient (F_{ROH}), genome coverage that resulted from the chosen parameter settings was 0.99, which allows ROH detection for 99% of the autosomal genome (Gorssen et al., 2021). Median F_{ROH} was 0.08 (Interquartile range = 0.06–0.18), with the highest F_{ROH} (0.20) estimated in H3 (Fig. 2). There were five animals with an estimated $F_{\text{ROH}} < 0.01$; however, four of these horses (H4-H7) had a call rate $\leq 86\%$. Segments of ROH ≥ 16 Mb were found in eight horses, with six of these horses coming from G1. Comparisons of ROH regions between individuals yielded one ROH island that was present in over 50% of the sampled horses. The ROH island was located on *equus caballus* chromosome (ECA) 8 and had a total length of 887 kb. There was no overlap detected between this region and any previously characterized QTL for reproduction and health traits.

4. Discussion

While it is not possible to make precise statements regarding relatedness (e.g. parent-offspring vs. full-siblings) of individuals due to a lack of birth records or pedigree information, by estimating the relationship coefficient it is possible to evaluate the degree of the relationship and hence gain insight into the sample population structure [19]. Since the estimation was carried out with two programs, they provide a range for r within the sampled horses [13]. Both programs yielded similar patterns and show two subgroups of horses (G1, G2), three horses highly related to the rest of the sample (H4, H5, H7), and two horses minimally related to the rest of the sample (H23, H24) (Fig. 1). The minimally related horses would be particularly valuable for any proposed domestic breeding program. While low individual call rates (< 86%) for four horses may have influenced estimates of r , estimates of r were based on 56,916 SNPs, thus any influence is expected to have been minimal [13]. Results from the current analyses also indicated that there are horses with F_{ROH} close to zero. However, it must be noted that 4/5 of these horses had call rates below what is generally considered as the minimum threshold for individual call rates (i.e. 95%) [11,20,21]. This increase in missing calls typically leads to segments not being classified as ROH

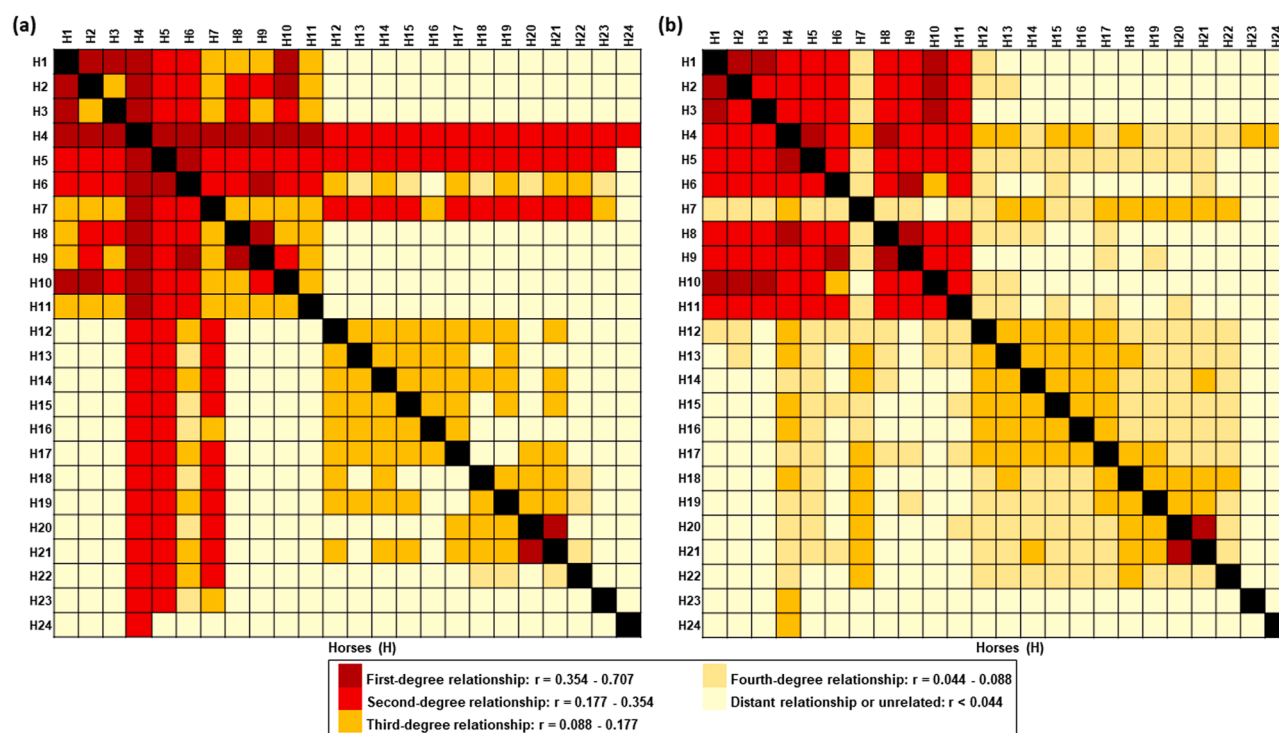


Fig. 1. Degree of relationships for 24 Timor ponies based on relationship coefficients (r) calculated in PLINK (a) and KING (b).

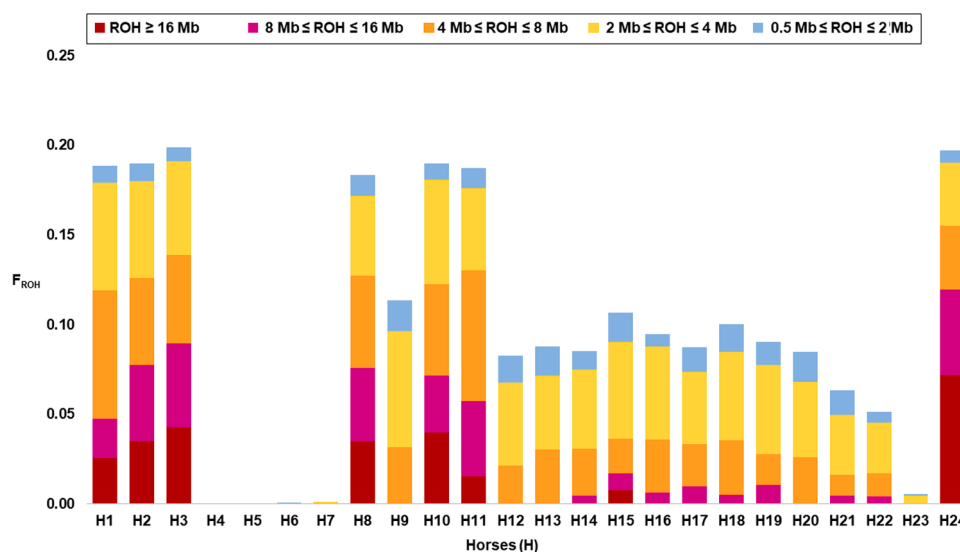


Fig. 2. Inbreeding coefficients (F_{ROH}) of 24 Timor ponies stratified by length (Mb) of ROH segments.

segments in PLINK as it is highly probable that these segments are interrupted by more than one missing call, placing them above the maximal number of missing calls allowed per window. Thus, horses with a reduced call rate may have an underestimated F_{ROH} . Although increasing the number of missing calls permitted within a window may, in part, help to resolve this, it can also artificially inflate F_{ROH} . Nevertheless, there are some horses (H21-H23) with individual call rates $> 95\%$ that have $F_{ROH} < 7.5\%$. Within these horses, no ROH segments greater than 16 Mb in length were identified, thus recent inbreeding events are unlikely [21]. Horses with ROH segments greater than 16 Mb in length were identified in eight horses and these individuals had F_{ROH} above the sample average. Moreover, 6/8 of these horses fell within G1 (Figs. 1 and 2), thus horses from the Cobourg Peninsula and their

descendants appear to have more recent inbreeding than the horses from G2.

5. Conclusion

Overall, lengths of the ROH segments indicate that recent inbreeding events are likely to only have occurred in a third of the horses sampled and the results illustrate that there are distantly related and minimally inbred horses within the TPs sampled. Additionally, there was no overlap between the ROH island on ECA 8 and any previously characterized QTLs for reproduction and health traits [18]. In general, these results are promising for the success of a domestic TP breeding program as there is clear evidence of genetic variation within the sampled TPs

and no obvious genomic regions of concern related to fertility and health. However, considering the low number of horses sampled, there would certainly still be substantial benefits to incorporating additional TPs, either from Australia or from Timor into a domestic breeding program.

Ethical Statement

The experimental protocols used in the current study were approved by the Ethics committees at each respective University in line with each university's animal ethics guidelines. All procedures were conducted in accordance with the guide for animal care and were performed by qualified staff.

Funding

This work was supported by a private contribution from Barb Bleicher.

CRediT authorship contribution statement

Doris E. Fröhlich: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **Barbara Wallner:** Writing – review & editing, Conceptualization. **Rytis Juras:** Writing – review & editing, Resources, Conceptualization. **E. Gus Cothran:** Writing – review & editing, Resources, Conceptualization. **Brandon D. Velie:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

None of the authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

References

- [1] Hendricks BL. International encyclopedia of horse breeds. 1st ed. Norman: University of Oklahoma Press; 1995.

- [2] Fijn N. Encountering the Horse: Initial Reactions of Aboriginal Australians to a Domesticated Animal. *Aust Humanit Rev* 2017;62:12–5.
- [3] Wesley P. Mounted police keep peace in Timor. *AFP News* 2001;100:12–3.
- [4] Haller M. Der neue Kosmos-Pferdeführer. 3rd ed. Stuttgart: Kosmos; 2003.
- [5] Bettencourt EMV, Tilman M, Narciso V, Silva Carvalho ML, Sousa Henriques PD. The Livestock Roles in the Wellbeing of Rural Communities of Timor-Leste. *Rev Econ Soc Rural* 2015;53:63–80.
- [6] Khanshour AM, Juras R, Cothran EG. Microsatellite analysis of genetic variability in Waler horses from Australia. *Aust J Zool* 2013;61:357–65.
- [7] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am Hum Genet* 2007;81:565–6.
- [8] Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W. Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010;26:2867–73.
- [9] Core Team R. R: A Language and Environment for Statistical Computing. *R Found Stat Comput* 2015. <http://www.R-project.org>. cited 20.10.2023.
- [10] Kamiński S, Hering DM, Jaworski Z, Zabolewicz T, Ruś A. Assessment of genomic inbreeding in Polish Konik horses. *Pol J Vet Sci* 2017;20:603–5.
- [11] Velie BD, Solé M, Jäderkvist Fegraeus K, Rosengren MK, Røed KH, Ihler C, et al. Genomic measures of inbreeding in the Norwegian–Swedish Coldblooded Trotter and their associations with known QTL for reproduction and health traits. *Genet Sel Evol* 2019;51.
- [12] Tinker NA, Mather DEKIN. Software for Computing Kinship Coefficients. *J Hered* 1993;84:238.
- [13] Ramstetter MD, Dyer TD, Lehman DM, Curran JE, Duggirala R, Blangero J, et al. Benchmarking Relatedness Inference Methods with Genome-Wide Data from Thousands of Relatives. *Genetics* 2017;207:75–82.
- [14] Meyermans R, Gorssen W, Buys N, Janssens S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *Bmc Genomics [Electronic Resource]* 2020;21.
- [15] Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. *BMC Genet* 2012;13.
- [16] Ferencaković M, Sölkner J, Curik I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet Sel Evol* 2013;45.
- [17] Gorssen W, Meyermans R, Janssens S, Buys N. A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. *Genet Sel Evol* 2021;53.
- [18] Hu ZL, Park CA, Reecy JM. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Res* 2021; 50:956–61.
- [19] Sun-Wei G. Proportion of Genome Shared Identical by Descent by Relatives: Concept, Computation, and Applications. *Am J Hum Genet* 1995;56:1468–76.
- [20] Hill EW, McGivney BA, MacHugh DE. Inbreeding depression and durability in the North American Thoroughbred horse. *Anim Genet* 2003;54:408–11.
- [21] Mastrangelo S, Tolone M, Di Gerlando R, Fontanesi L, Sardina MT, Portolano B. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal* 2016;10:746–54.