A PHYLOGENETIC ANALYSIS OF BRUMBIES

F.W. Nicholas[1], E.G. Cothran[2], L. Jermiin[3], and B. Nesbitt[4]
[1]Reprogen, Faculty of Veterinary Science, University of Sydney, NSW 2006, Australia
[2]Department of Veterinary Science, University of Kentucky, Lexington, KY 40546, USA
[3]School of Biological Sciences, University of Sydney, NSW 2006, Australia
[4]NSW National Parks and Wildlife Service, PO Box 170 Dorrigo 2453

SUMMARY
A brumby is a "wild" horse that is the descendant of domesticated horses that escaped from, or were released from, grazing properties. "Wild" horses in the Guy Fawkes River National Park (GFRNP) in New South Wales have relatively high genetic similarity with Arabian-type breeds and/or saddle and harness light horses (for example, Thoroughbreds), and are also genetically similar to Walers (horses bred in the colony of New South Wales for the Army remount trade in India, South Africa and the Middle East). Horses in the GFRNP have a relatively low level of inbreeding, and are not a significant reservoir of unique genes. All of these attributes are consistent with local anecdotal evidence (which suggests a more-or-less continual introduction of "outside" blood). The genetic evidence is, therefore, consistent with the concept of brumby.
Keywords: brumby, phylogeny, horse

INTRODUCTION
The Guy Fawkes River National Park (GFRNP) comprises more than 62,000 hectares approximately 100 km north-east of Armidale, NSW. As in many other conservation areas throughout the world, the GFRNP is home not only to a wide range of native flora and fauna, but also to non-native animals, most commonly descendants of domesticated animals that escaped or were released into the area. In Australia, the term "brumby" is applied to horses that exist under such circumstances. In 2001, in the context of considering the humane removal and future management of the horses in the GFRNP, the NSW Minister for the Environment established a Working Party to investigate the heritage value of horses. As part of that investigation, blood samples were obtained from, and conformation was assessed on, 16 horses that had been captured and removed from the GFRNP. The aim of the sampling and assessing was to test the claim that horses in the GFRNP have sufficient genetic uniqueness to warrant them being conserved on that basis. This paper describes the results of phylogenetic analyses conducted on genotypic data obtained by genotyping the 16 horses at a standard set of loci, and on the conformation assessments.

MATERIALS AND METHODS
Blood group systems and other traditional loci. Blood samples were obtained from 16 horses from the GFRNP (Guy Fawkes horses) and from 20 Walers located on properties scattered throughout Australia. (Walers are horses bred in the colony of New South Wales, for the Army remount trade in India, South Africa and the Middle East.) The samples were sent to the Australian Equine Genetics Research Centre at the University of Queensland, where they were processed and genotyped at seven standard blood-group systems (A, C, D, K, P, Q, U) and nine standard electrophoretic loci (albumin, transferrin, esterase, alpha-1-beta-glycoprotein, protease inhibitor, phosphogluconate dehydrogenase,

phosphohexose isomerase, haemoglobin, vitamin D binding protein) under the guidance of Dr Helen Arthur, following international standard protocols (as cited, for example, by Cothran *et al.* 2001). These data were then added to the extensive database of genotyping results from more than 50 wild-horse populations and more than 100 recognised breeds, maintained by one of us (EGC) at the University of Kentucky.

Following the standard procedure in the Kentucky laboratory (e.g. as in Cothran *et al*., 2001), the following parameters were estimated from the nine electrophoretic loci: observed heterozygosity (*Ho*), Hardy-Weinberg expected heterozygosity (*He*), population inbreeding ($F_{IS} = 1 - Ho/He$). In addition, the following parameters were estimated from all 16 loci: Hardy-Weinberg expected heterozygosity (*Het*), effective number of alleles (*Ae*), and total number of variants (*TNV*). The reason for two separate estimates of expected heterozygosity is that, because heterozygotes are not always distinguishable at blood-group loci, it is not possible to estimate observed heterozygosity at these loci. The first estimate of expected heterozygosity (*He*) was calculated for direct comparison with the estimate of observed heterozygosity obtained from the electrophoretic loci; the second estimate of expected heterozygosity applies to all 16 loci. A restricted-maximum-likelihood (RML) phylogenetic analysis was also conducted on the allele-frequency data, using the PHYLIP package (Felsenstein, 1989).

Conformation data. The collection and analysis of data were based on the procedures detailed by Jordana *et al.* (1995), who provided a mainly subjective scoring system for each of 30 mainly morphological traits (their Table 1). In their own analysis, Jordana *et al.* scored an "ideal specimen" of each of 20 horse breeds from Western Europe, North Africa and South America, and two extinct wild-horse breeds. Their raw data (presented in their Table 2) thus comprised a table of scores with 30 columns and 22 rows. For the present analysis, 14 of the 16 Guy Fawkes horses that were blood sampled were also scored according to the Jordana *et al.* system, with the same three people (two of whom were experienced horse people) collaborating in the scoring of all 14 horses. In addition, all 20 of the Walers that were blood sampled were also scored according to the Jordana *et al.* system. However, since the Walers were from geographic locations scattered around Australia, it was not possible to have all horses scored by the same persons. Instead, experienced horse people at each site scored the horses that were sampled at that site. Thus, variation in the Waler data includes variation between scorers, whereas the Guy Fawkes data do not. In order to add these data to that already provided by Jordana *et al.*, it was necessary to reduce the Australian data to a single row of scores for each population. This was achieved by calculating a simple average of the scores for each trait within each population, and then expressing each average to the nearest whole number. Finally, because of its importance as a horse breed and because of its likely relationship to Guy Fawkes horses and Walers, a row of scores for an "ideal specimen" of an Australian Thoroughbred was provided by an experienced equine veterinarian, Dr Paul McGreevy. After adding the three rows of Australian data to the data of Jordana *et al.*, the raw data for the present analyses compromised 30 columns and 25 rows of data.

For a phylogenetic analysis, version 4.0b8 of PAUP* (Swafford, 1998) was used in a manner very similar to that used by Jordana *et al.* Specifically, a heuristic search was conducted using the

maximum-parsimony criterion, with the following settings: best trees only, random addition of breeds, 100 runs, swapping algorithm, tree-bisection-reconnection (TBR), traits unordered.

RESULTS

Table 1 shows the parameters estimated from the Guy Fawkes horses and from Walers, in comparison with eight recognised breeds, and with averages for wild-horse populations and all domestic breeds. Except for the Guy Fawkes and Waler horses, the data are not from Australian horses: instead, they come from ECG's extensive database of primarily US data.

Table 1 Estimates of parameters from Guy Fawkes horses and from Walers, in comparison with estimates from other breeds and populations

Population	***N***	***Ho***	***He***	F_{IS}	***Het***	***Ae***	***TNV***
Waler	20	0.365	0.392	0.069	0.447	2.416	62
Guy Fawkes	15	0.273	0.286	0.045	0.338	2.173	49
Thoroughbred	265	0.294	0.288	-0.019	0.325	2.009	64
Arabian	117	0.307	0.327	0.061	0.376	2.132	67
Andalusian	140	0.348	0.362	0.039	0.425	2.508	75
Shetland Pony	50	0.368	0.407	0.095	0.452	2.595	71
Welsh Pony	42	0.388	0.387	-0.002	0.453	2.603	76
American Saddlebred	259	0.404	0.409	0.013	0.435	2.625	96
Peruvian Paso	141	0.451	0.445	-0.014	0.469	2.761	77
Belgian Draft	82	0.427	0.415	-0.028	0.451	2.386	66
US Wild Horse Mean	54[a]	0.360	0.351	-0.035	0.385	2.218	53.5
Standard Deviation		0.051	0.053	0.118	0.067	0.339	12.5
Domestic Horse Mean	118[a]	0.371	0.365	-0.014	0.414	2.398	65.4
Standard Deviation		0.049	0.043	0.065	0.039	0.253	11.1

a. number of populations/breeds

It can be seen that all three heterozygosities (*Ho*, *He* and *Het*) of Guy Fawkes horses are somewhat lower than the average for wild-horse populations and for domestic breeds. Using the measure of heterozygosity estimated from all 16 loci (*Het*), a useful way to summarise these results is to say that Guy Fawkes horses have 12% less genetic variability than the average of wild-horse populations, and 18% less genetic variability than the average of recognised domestic breeds. Viewed from another perspective, the genetic variability in Guy Fawkes horses is 0.70 of a standard deviation below the average of wild-horse populations, and nearly two standard deviations below the average of recognised domestic breeds. In contrast, Walers have 8% more genetic variability (0.85 of a standard deviation higher) than the average of recognised domestic breeds.

Of the 49 different variants (alleles) detected in Guy Fawkes horses, 47 occur in recognised breeds in Australia (Helen Arthur, personal communication); and all 62 genetic variants (alleles) detected in

Walers occur in recognised breeds in Australia (Helen Arthur, personal communication). In other words, neither Guy Fawkes horses nor Walers are genetically unique to any significant extent.

In the present context, F_{IS} is a measure of the extent of inbreeding in the populations from which the samples were drawn. The relatively low values of 5% and 7% for Guy Fawkes horses and for Walers, respectively, indicate that neither population is inbred to any extent. This is consistent with continued migration into the populations, which is suggested by local anecdotal evidence.

The consensus phylogenetic tree from the RML analysis is shown in Figure 1a. This tree is a majority-rule, strict-consensus tree from 50 separate RML trees. The numbers atthe forks indicate the number of times the groups consisting of the populations to the right of that fork occurred among the total of 50 trees that were constructed. In other words, the closer the number is to 50, the greater the confidence in concluding that the populations to the right are more similar to each other than to other populations.

It is evident from Figure 1a that Guy Fawkes horses are most genetically similar to Arabian-type breeds, and next-most-genetically-similar to saddle and harness light horses. The opposite applies for Walers: they are most genetically similar to saddle and harness light horses, and next-most-genetically-similar to Arabian-type breeds. The two groups to which the Guy Fawkes Horses and Walers belong cluster next to each other, indicating a greater genetic similarity between these two groups than between either of these two groups and any other group of breeds.

Conformation. Figure 1b shows the sole maximum-parsimony tree that resulted from the heuristic search, rooted with the Tarpan and Przewalski's horse as outgroups. The branch lengths are indicative of the number of trait-stage changes (roughly: the number of steps or changes in trait) required to move from one population to the next. These numbers are shown on thebranches. The Arabian-type breeds and the saddle and harness light horses are all clustered together, and, consistent with the genetic results, both Guy Fawkes horses and Walers are in this cluster.

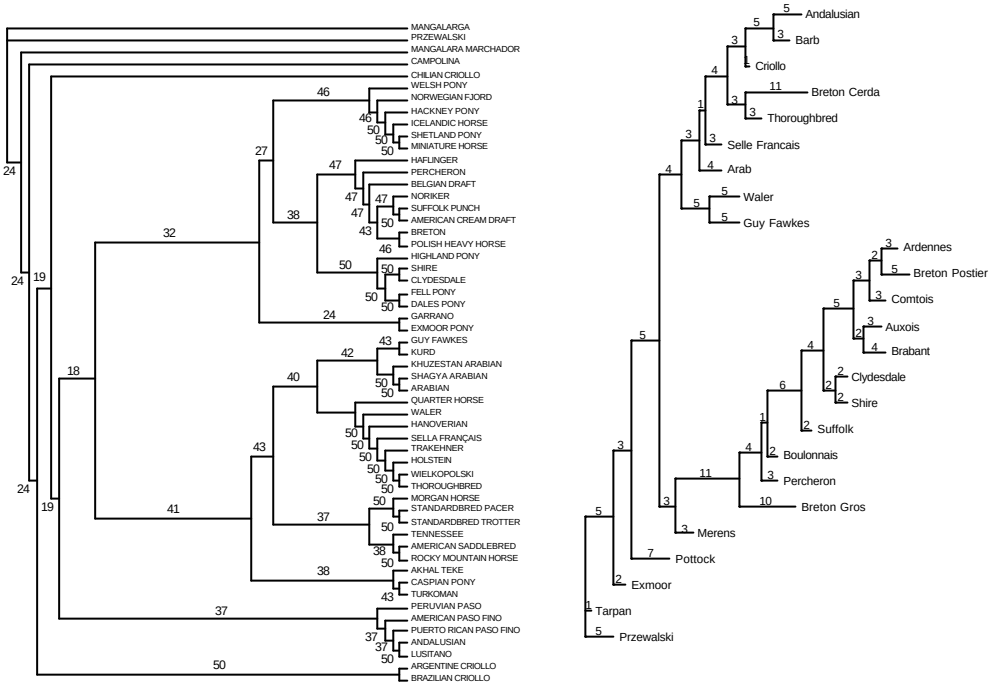

Figure 1. (a; left) Phylogenetic tree based on blood-typing data, showing that Guy Fawkes horses and Walers are most closely related to Arabian-type breeds and saddle and harness light breeds. The tree represents the consensus from 50 replicate trees that were generated from the data. The numbers at the forks indicate the number of times the groups consisting of the populations to the right of that fork occurred among the total of 50 trees that were constructed. In other words, the closer the number is to 50, the greater the confidence in concluding that the populations to the right are more similar to each other than to other populations. (b: right) Phylogenetic tree based on conformation data, showing Guy Fawkes horses and Walers in the same group as Arabian-type breeds and saddle and harness light breeds. The number on each branch of the tree and the length of each branch indicate the extent to which populations differ.

DISCUSSION

The comparison would have been more powerful if the Guy Fawkes horses and Walers could have been compared with Australian samples of the recognized breeds. Because of various constraints, it was not possible to include such data in the analyses reported here. However, the authors have compared Australian and US samples of the major breeds (unpublished results), and it is evident that samples from the two countries are very similar. We can therefore be confident that the same conclusions would have been drawn if the comparisons had been made against Australian samples of

established breeds. Although based on very different data from the genetic evidence above, the results from the conformation data are remarkable similar to those from the genetic data, except that there are fewer clusters, because this form of analysis is less powerful than the genetic analysis.

ACKNOWLEDGEMENTS

Sincere thanks to Graeme Baldwin and Velda Chaplin, members of the Working Party who for their roles in organising the blood sampling and scoring of conformation. The authors are especially grateful to the many people who assisted in these operations, especially Erica Baldwin, Monika Darke, and the owners of the Guy Fawkes horses and Walers that were kindly made available for this study. Thanks are also due to Drs Kevin Bell and Helen Arthur and to the Australian Stud Book for enabling the blood samples to be processed and genotyped. This paper has been extracted from information included in volumes 1 and 2 of the report of the Working Party (Anon., 2002).

REFERENCES

Anon. (2002) http://www.npws.nsw.gov.au/history/guy_fawkes_horses/index.html (15 February 2003)

Cothran, E.G., van Dyk, E. and van der Merwe, F.J. (2001) *J. Sth African Vet. Assoc.* **72**: 18.

Felsenstein, J. (1989) *Cladistics* **5**: 164.

Jordana, J., Pares, P.M. and Sanchez, A. (1995) *J. Equine Vet. Sci.* **15**: 320.

Swafford, D.L. (1998) PAUP*. Phylogenetics analysis using parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, Mass.