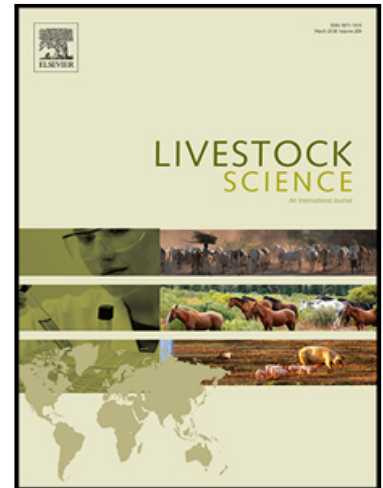


Journal Pre-proof

Pedigree analysis of the racing line Quarter Horse: Genetic diversity and most influential ancestors

Ricardo António da Silva Faria , António Pedro Andrade Vicente ,
Alejandra Maria Toro Ospina ,
Josineudson Augusto II Vasconcelos Silva

PII: S1871-1413(21)00092-5
DOI: <https://doi.org/10.1016/j.livsci.2021.104484>
Reference: LIVSCI 104484



To appear in: *Livestock Science*

Received date: 13 June 2020
Revised date: 26 February 2021
Accepted date: 18 March 2021

Please cite this article as: Ricardo António da Silva Faria , António Pedro Andrade Vicente , Alejandra Maria Toro Ospina , Josineudson Augusto II Vasconcelos Silva , Pedigree analysis of the racing line Quarter Horse: Genetic diversity and most influential ancestors, *Livestock Science* (2021), doi: <https://doi.org/10.1016/j.livsci.2021.104484>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Published by Elsevier B.V.

Highlights

- Pedigree data of the racing line Quarter Horse since 1747.
- Failures in the breeding and conservation programs of the racing line.
- Loss of genetic diversity and concentration in a reduced number of ancestors.
- Provided the names of the most influent ancestors.
- The results allow guidelines for breeders of the racing line Quarter Horse.

Pedigree analysis of the racing line Quarter Horse: Genetic diversity and most influential ancestors

Authors: Ricardo António da Silva Faria^{a,b*}, António Pedro Andrade Vicente^{c,d}, Alejandra Maria Toro Ospina^a, Josineudson Augusto II Vasconcelos Silva^e

^aUniversidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, CEP 14.884-900, Jaboticabal, São Paulo, Brazil.

^bEscola Profissional Agrícola Quinta da Lageosa, Apartado 32, 6254-909 Belmonte, Portugal.

^cEscola Superior Agrária do Instituto Politécnico de Santarém, Apartado 310, 2001-904 Santarém, Portugal.

^dCIISA - Faculdade de Medicina Veterinária, Universidade de Lisboa, 1300-477 Lisboa, Portugal.

^eUniversidade Estadual Paulista (Unesp), Faculdade de Medicina Veterinária e Zootecnia, CEP 18.618-307 - Botucatu, São Paulo, Brazil.

Corresponding author: Ricardo António Silva Faria

***E-mail:** fariasky@gmail.com

Postal Address: FMVZ - Unesp, DMNA - Fazenda Experimental Lageado

Rua José Barbosa de Barros, nº 1780 CEP: 18.618-307 - Botucatu - SP - Brasil

E-mails of the co-authors in the same order of publication:

Antônio Pedro Andrade Vicente: apavicente@gmail.com

Alejandra Maria Toro Ospina: toroospina92@gmail.com

Josineudson Augusto II Vasconcelos Silva: agosto.ii@unesp.br

Journal Pre-proof

Abstract – The aims of this study were to evaluate the genetic diversity of the racing line Quarter Horse breed in Brazil and to provide data on the most influential ancestors by pedigree analysis. All horses that had participated in sprint races in Brazil between 1978 and 2016 were evaluated. There were 5,861 athletes horses (2,474 males and 3,387 females) born between 1971 and 2014. These animals were referred to as the racing population (Pr), which was divided into three subpopulations to evaluate three complete decades (84.5% of Pr). These subpopulations consisted of 1,712, 1,604 and 1,461 animals born in the 1980s (sP80), 1990s (sP90) and 2000s (sP00), respectively. The quality of the pedigree was assessed based on the number of complete equivalent generation traced. There were 5.4 (Pr), 4.7 (sP80), 5.7 (sP90) and 6.3 (sP00) generations, which permitted accuracy of the conclusions reported. The generation intervals were long, with 13.2 (sire-son/daughter) and 10.6 (dams-son/daughter) years (Pr). The coefficients of inbreeding and average relatedness were 0.95% and 1.84% for Pr and 1.60% and 2.56% for sP00, respectively. In Pr and sP00, the effective population size (N_e) based on ΔF_i was 215 and 144, respectively, with 68 and 41 animals per generation. The probability of gene origin given by the effective number of founders (f_e), ancestors (f_a) and founder genome equivalents (f_{ge}) was 192, 61 and 34, respectively, for Pr, with considerable losses of genetic diversity over the last decades evaluated. The number of ancestors (founders or not) that explained the total genetic diversity was 1,587. This number is considered high; however, only 32 and 9 ancestors explained 50% of the genetic diversity of Pr and sP00, respectively. The 10 most influential ancestors of sP00 explained 52.2% of the total genetic diversity. The loss of genetic diversity and concentration in a reduced number of breeding animals indicate failures in the breeding and conservation programs of the racing line. Mating designs in conjunction with the selection of breeding animals should be promoted and goals should be defined for racing line of the Quarter Horse breed in Brazil.

Keywords: effective population size; generation interval; inbreeding; pedigree chart

1. Introduction

Quarter Horses (QH) are the fastest horses in the world in short-distance races (Nielsen et al., 2006) and the most versatile (Petersen et al., 2014). In addition, this breed represents the largest population of registered horses in Brazil (514,316 animals) (ABQM, 2019) and in the world (2,906,070 animals) (AQHA, 2015). The QH breed comprises three groups (lines) of genetically distinct animals (Petersen et al., 2014), which are defined based on their performance: stock horses (working livestock and Western competitions), halter/pleasure horses, and racing horses.

Studies evaluating the genetic diversity and most influential ancestors of all animals of the QH breed (considering the three lines) based on genealogical data were conducted in the United States (Tunnell et al., 1983) and Brazil (Faria et al., 2018a). However, no emphasis was given considering each line separately, as reported by Marchiori et al. (2019) and Petersen et al. (2014).

The first official sprint races involving QH animals were carried out in 1941 in the United States and in 1978 in Brazil. The total prize money distributed from 1978 to 2016 in Brazil was US\$43,887,504.9, with an average prize money per animal/racing of US\$13,976.9 (values updated in June 2019) (Faria and Silva, 2019). These values are expressive and indicate a profitable horse industry.

The monitoring and evaluation of genetic diversity (Maignel et al., 1996) and of most influential ancestors (Boichard et al., 1997) permit to design and develop genetic breeding programs for athletes horses aimed at preserving genetic diversity. Studies that evaluate the population structure and identify its important ancestors considering only the racing line of the QH breed are not available but are important.

The aims of this study were to evaluate the evolution of genetic diversity over time in the racing line of the QH breed in Brazil and to describe its most influential ancestors.

2. Material and Methods

2.1. Animals, (sub)populations, pedigree quality and generation interval

The pedigree database of the racing QH line was provided by the Sorocaba Jockey Club (JCS) representing 61.6% of the total of pedigree information (5.861 athletic horses and their ancestors - 624 sires and 2.550 dams) and from Brazilian

Quarter Horse Breeders Association (ABQM) with 35.9% of the total of pedigree (4.066 ancestors).

Ancestors of animals not known in the file that were needed to increase the quality of the pedigree were added. The data of the 324 ancestors (2.8% total of pedigree) were obtained from the internet: Pedigree Online's All Breed Database (www.allbreedpedigree.com). The database contained information about the name of the animal, sire, dam, sex and date of birth. From the 5,861 athletes horses with racing records in Brazil (racing population, Pr) born between 1971 and 2014, the pedigree file was constructed with the ascendants and included 11,425 animals born since 1747.

Using the Pr file (5,861 animals), three subpopulations (total of 4,777 animals) representing the following decades were evaluated: 1980-1989 (sP80), 1990-1999 (sP90) and 2000-2009 (sP00), with 1,712, 1,604 and 1,461 animals, respectively. The subpopulations allow the analysis of animals that contributed or may have contribute to subsequent generations (Siderits et al., 2013).

The quality of the pedigree is represented by the mean number of complete equivalent generation traced (EG) and was calculated as described by Maignel et al. (1996). The generation interval (GI) was calculated considering the birth of parents when their offspring reproduced (James, 1972) for two different paths: sire-son/daughter and dam-son/daughter.

2.2. Genetic diversity

2.2.1. Inbreeding parameters

The individual inbreeding coefficient (F) was calculated according to Meuwissen and Luo (1992) and the individual average coefficient of relatedness (AR) was obtained as described by Gutiérrez and Goyache (2005). The individual increase in inbreeding (ΔF_i) was calculated according to González-Recio et al. (2007) and the effective population size (N_e) based on ΔF_i was estimated as proposed by (Gutiérrez et al., 2009). The latter parameter was used to calculate genetic drift. Additionally, N_e was calculated based on the increase in inbreeding per generation (Gutiérrez et al., 2008).

The genetic structure of Pr and of the subpopulations was evaluated using Wright's F statistics (Wright, 1965), calculated as described by Caballero and Toro (2000, 2002), which allows the evaluation of differences between the generated files. The individual inbreeding coefficient in relation to the total population (F_{IT}), the average inbreeding coefficient of the subpopulation in relation to the total population (F_{ST}), and the inbreeding coefficient of the individual in relation to the subpopulation (F_{IS}) were calculated.

2.2.2. Number of founders, probability of gene origin and genetic drift

Ancestors with no known parent are considered as founders, indicated by number of founders (N_f). The probability of gene origin was evaluated based on the effective number of founders (f_e) as described by Lacy (1989), effective number of ancestors (f_a) according to Boichard et al. (1997), and effective number of founder genome equivalents (f_{ge}) calculated according to Caballero and Toro (2000).

Genetic drift is defined as a random change in the allele frequencies of the population, which occurs at a rapid pace when the population undergoes a drastic reduction in its effective size (Frankham, 1996). The unequal contribution of ancestral founders, the presence of a bottleneck effect and the loss of founder alleles among generations in the populations are observed when the f_e/N_f (number of founders), f_e/f_a and f_{ge}/f_e ratios, respectively, are different and distant from 1 (Boichard et al., 1997).

2.2.3. Genetic contributions of most influential ancestors and pedigree chart

The most influential ancestors (founders or not) that explained the total genetic diversity of the populations and their cumulative marginal genetic contributions were calculated as described by Boichard et al. (1997). The data of the 10 most influential ancestors of Pr and sP00 are reported, as well as their cumulative marginal genetic contributions. The pedigree chart of the most influential ancestors was constructed to allow visual observation of the relationship between animals.

The SPSS25 (IBM, 2017) and PedGraph (Garbe and Da, 2008) programs were used for data preparation, statistical analysis of the pedigree, and construction of the pedigree chart. The genetic structure parameters were obtained with the ENDOG v4.8 program (Gutiérrez and Goyache 2005).

3. Results

3.1. Animals, (sub)populations, pedigree quality and generation interval

The number of horses registered in official races in Brazil decreased in the last decades by 6.3% and 8.9% from sP80 to sP90 and from sP90 to sP00, respectively (Table 1). The distribution by sex indicates a higher percentage of females, but this number tended to decrease in the last three decades (Table 1).

A number of EG larger than five indicates the quality of the pedigree information. In this study, the largest number (6.3 generations) was observed in the last generation (sP00). The GIs were long, > 10 years for both paths evaluated (sire-son/daughter and dam-son/daughter). There was a decrease in the GI in sP90, while a GI longer than the overall mean was observed in sP00. Mares had shorter GIs than stallions (Table 1).

3.2. Genetic diversity

3.2.1. Inbreeding parameters

The F coefficient increased in the last decades, with a higher overall mean in sP00. The same trend was observed for AR and ΔF_i (Table 1). The difference between the values of these parameters considering all animals evaluated and only the animals with inbreeding coefficient different from zero ($F \neq 0$) was small (< 1%). The number of animals with $F \neq 0$ increased considerably in the last two decades studied, from 36.6% in the 1980s (sP80) to 70.4% in sP90 and 88.8% in the last decade (sP00). However, the level of inbreeding remained low (Table 1).

The N_e values calculated based on ΔF_i were higher than the N_e per generation (Table 1). The highest value was observed in sP80 and the lowest in sP00. The N_e per generation in sP00 was lower than the limit recommended by the FAO/ONU (FAO, 1998). Analysis of genetic differences using Wright's F statistics provided negative values for F_{IS} and F_{IT} , with the F_{IT} values being close to zero. The F_{ST} values were positive and the same in all subpopulations (Table 1).

3.2.2. Number of founders, probability of gene origin and genetic drift

The N_f was high in Pr, with the observation of considerable loss over the last three decades. The probabilities of gene origin given by f_e , f_a and f_{ge} also decreased. The lowest values were observed in sP00 (Table 2).

The genetic drift calculated from the probabilities of gene origin and N_f using the f_e/N_f , f_a/f_e and f_{ge}/f_e ratios provided results distant from 1 (ideal value). Values close to unity were observed in sP80 (Table 2).

3.2.3. Genetic contributions of the most influential ancestors and pedigree chart

The number of ancestors (founders or not) explaining total genetic diversity (100%) in the Pr, sP80, sP90 and sP00 racing groups corresponded to 69.9%, 50.1%, 46.4% and 42.0% of total N_f , respectively (Table 2 and 3). In the last decade (sP00), with only 1.9% of the ancestors explained 50% of the genetic diversity of racing line QH in Brazil (Table 3). The cumulative marginal genetic contributions of the 10 most influential ancestors (Pr) were high (Table 3) and increased considerably in the last decade evaluated (sP00). Of the 10 ancestors that most influence the total genetics of QH racehorses in Brazil, only 4 and 8 ancestors (Table 3) appeared in the 1980-89 (sP80) and 1990-99 (sP90) decades. Two changes were observed among the 10 most influential ancestors of sP00 compared to Pr, with the entry of Apollo VM and Little Smoothie and exit of Top Deck and Lady's Moon (two ancestors who have decreased their contributions over the decades). The three most influential ancestors increased their genetic contribution by 4.06% (Dash for Cash), 2.92% (Easy Jet), and 2.79% (Beduino) in sP00, when compared with Pr (Table 3).

Evaluating the Thoroughbred or Thoroughbred half-blooded ancestors, included in the 10 main ancestors of the breed (Table 3), we observed that the marginal genetic contributions increased 6.49% from sP80 to sP90 and 1.62% from sP90 to sP00.

The total number of offspring of the 10 most influential ancestors was 1,125 animals (Table 3). In this group, 75.9% were athletes sprint racing horse in Brazil, progeny of the following ancestors (number of athletes offspring): Signed to Fly (342), Apollo VM (178), Blazen Bryan (156), Lady's Moon (136), Streakin Six (20), Dash for Cash (9), Easy Jet (3) and Little Smoothie (1). The average pedigree quality

(EG) of the most influential ancestors was 5.0 generations (Table 3), ranging from 3.5 to 6.1 generations. The lowest numbers of EG and the only ones < 5.0 generations were observed for the three Thoroughbred animals (Three Bars, Beduino, and Top Deck).

Analysis of the relatedness coefficients between the 10 most influential ancestors of Pr and the 10 ancestors of sP00 identified 12 ancestors (8 of them being the same animals in Pr and sP00). Of the 66 possible individual results of AR among the 12 ancestors, 25 AR coefficients were higher than 5.0%, with a maximum value of 31.35% (Table 4).

Creation of a pedigree chart (Supplemental Figure S1) permitted to verify the relationship of the most influential ancestors of Pr and sP00 (Table 4). The description including only additional 39 ancestors (Supplemental Figure S1), offspring that became breeding animals, demonstrates the genetic link between the most influential ancestors.

4. Discussion

4.1. Animals, (sub)populations, pedigree quality and generation interval

A higher percentage of females in the races has also been described for Thoroughbreds (Mota et al., 1998; Mota et al., 2005) and Arabian horses (Ekiz and Kocak, 2005), as well as for the QH breed (Buttram et al., 1988; Corrêa e Mota, 2007; Silva et al., 2014). A smaller number of males with race records is not the result of a smaller number of births but is rather due to the fact that not all males born are registered and perform racing activities. This choice of breeders is related to the less permissible selection criteria of future stallions, when compared to females. In females, the lack of selection criteria and the need of more females for reproduction favors registration, in addition to the use of females as embryo recipients, an activity performed in Brazil since 1987 (Fleury et al., 1989). The higher percentage of males in sP00 is possibly the result of online registration and subsequent submission of the confirmation of paternity, reducing costs.

The pedigree quality data (EG) showed good values (Table 1), ensuring the accuracy of the present results. Studies in the literature reported smaller numbers of EG for the Mangalarga Marchador and Crioulo breeds of 2.38 (Baena et al., 2020)

and 3.51 (Maciel et al., 2014), respectively. On the other hand, a similar number of EG was obtained for purebred Lusitano and Maremmano of 5.7 generations for both breeds (Faria et al., 2018b; Giontella et al., 2019). Higher numbers of EG were reported for the Old Kladruber and Lipizzan breeds, with 15.1 (Vostrá-Vydrová et al., 2016) and 10.3 generations (Pjontek et al., 2012), respectively. The quality of the pedigree is important for determination of the results since it influences the effect of inbreeding depression, genetic bottlenecks, and genetic diversity. The data quality of the pedigree should be improved on the part of ABQM and the ancestors that exist knowledge on the part of AQHA should be inserted.

Long GIs are common in horses. For the QH breed, GIs of 10.5 years (all animals) and 9.9 years (racing line) were reported in the United States (Petersen et al., 2014) and of 10.8 years (all animals) in Brazil (Faria et al., 2018a). Taveira et al. (2004) obtained a GI of 7.1 years for the racing line of Thoroughbred horses. The long GI, especially in stallions (Table 1), leads to lower genetic gain. The difference in GI between stallions and mares (Table 1), as well as the longer permanence of males, is due to post-mortem sperm storage. The high standard deviations of GI (Table 1) suggest that some breeders use younger animals for reproduction. The GI is monitored in breeding programs in an attempt to reduce the interval and to obtain higher genetic gains. This goal can be achieved by using younger breeding animals and keeping them for a shorter period of reproduction.

4.2. Genetic diversity

4.2.1. Inbreeding parameters

The values of the inbreeding parameters (Table 1) were low and similar to those reported in the literature for QH animals. In the United States, Tunnell et al. (1983) described F coefficients ranging from 1.3% to 2.6% and Petersen et al. (2014) reported a coefficient of 1.6% as observed here for sP00. In Brazil, considering all QH lines, Faria et al. (2018a) reported an F of 1.1% to 1.6% and an AR of 1.0% to 1.2%.

The fact that QH breeders maintain low F and AR coefficients is not due to intentional control but to the large number of animals. This statement is possible based on the high percentage of animals with $F \neq 0$ (88.8% in sP00) and the F

coefficient has a value greater than half of the AR value (Table 1). The increase in F and in the number of animals with $F \neq 0$ in the last generation studied (sP00) suggests the possibility that all animals within one decade are related, aggravating inbreeding.

In a combined analysis of all QH lines in Brazil, Faria et al. (2018a) found an N_e based on ΔF_i of 167. Although a similar value was obtained in the present study (Table 1), a reduction in N_e was observed in the racing line, indicating a different genetic structure compared to all QH animals. This finding can be explained by the transition of different genes in each line, with lower gene flow in the racing population and consequently lower genetic diversity. This result are in accordance with the genome study by Marchiori et al. (2019), which reported N_e values of 50 and 60 for two lines, racing and cutting, respectively.

The N_e per generation of less than 50 in sP00 (Table 1) indicates the loss of genetic diversity. According to the FAO/ONU (FAO, 1998), an N_e less than 50 exceeds the minimum limit necessary for avoiding inbreeding depression. Studies on Italian Heavy Draught (Mancin et al., 2020) and Brazilian Mangalarga Marchador (Baena et al., 2020), reported N_e values of 97 and 1,471, respectively. However, differences between the population evaluated and other breeds reflect the number of animals, pedigree data, and probabilities of gene origin. Regular monitoring by ABQM and its breeders is necessary to avoid increases in the inbreeding parameters that directly influence N_e , compromising the conservation of genetic diversity.

The negative F_{IS} and F_{IT} values (Table 1) demonstrate the use of inbred mating considering the pedigree as part of racing animals. The decrease in F_{IS} over the last three decades confirms the increase in the number of animals with $F \neq 0$ (Table 1), indicating reduced awareness of breeders. The excessive use of certain breeding animals is visible considering the small difference between F_{ST} values in the subpopulations (Table 1), which suggests that genes of common ancestors are shared among the subpopulations (Cervantes et al., 2008). The lower frequency of shared genes between animals of sP90 and sP00 indicates a decrease in the number of breeding animals, mainly because of the intensification of the use of reproductive biotechnologies.

4.2.2. Number of founders, probability of gene origin and genetic drift

In the present study, the decreasing values of N_f , f_e , f_a and f_{ge} in all subpopulations (Table 2) indicate the predominant use of certain lineages of ancestors (founders or not). The difference between N_f and f_e suggests different contributions of founders to Pr and to the subpopulations evaluated. The f_a values indicate that the reduced number of ancestors is responsible for a large proportion of genetic diversity, with marked losses in sP00 (Table 2). The reduced f_{ge} values indicate a high loss of founder alleles.

The probabilities of gene origin obtained in the present study (Table 2) are closer to those of the Thoroughbred population (Schubertová et al., 2014), with f_e , f_a and f_{ge} of 949, 202 and 67, respectively, compared to the QH population including all lines (Faria et al., 2018a). For a complete evaluation of the genetic structure (racing line of the QH breed in Brazil), a genomic study similar to that performed by Petersen et al. (2014), should be conducted, allowing the identification of similar genetic regions between the two breeds.

A reduction in the N_e (Table 1) is seen in the last decades (subpopulations studied), with an effect on genetic drift and the consequent loss of founder alleles, reduced genetic variability and unequal contributions of founders (Table 2). A genetic drift (f_e/N_f) similar to the present study was described for Old Kladruber (Vostrá-Vydrová et al., 2016), Turkish Arab (Duru, 2017) and American Shire (Stephens and Splan, 2013) horses, with values of 0.09, 0.11 and 0.12, respectively. These results indicate the lack of conservation of genetic origin, with losses of connection and unequal contributions of founders since a value close to 1 would be ideal. Losses of founder alleles among generations (f_{ge}/f_e) were observed in the present study and in Old Kladruber (Vostrá-Vydrová et al., 2016) and QH (Faria et al., 2018a). Genetic bottleneck effects (f_e/f_a) were also reported for other breeds such as QH (Faria et al., 2018a) and American Shire (Stephens and Splan, 2013), indicating the use of a reduced number of breeding animals.

4.2.3. Genetic contributions of the most influential ancestors and pedigree chart

Analysis of the cumulative marginal genetic contributions of the most influential ancestors (Table 2) showed that a small number of animals were

responsible for the genetic diversity of the racing population. The six most influential ancestors are Dash for Cash, Easy Jet, Beduino, Three Bars, Signed to Fly and Top Deck, which were also identified in the study investigating all lines of the breed (Faria et al., 2018a), indicating their influence on different purposes, i.e., these stallions apparently transferred their genes to offspring in the different lines of the QH breed.

The ancestors Three Bars and Top Deck appear in the study of Tunnell et al. (1983) as athlete horses in sprint races and are among the animals with the greatest contribution to the population. The stallion Dash for Cash appears in the racing line in the study of Petersen et al. (2014) as being responsible for 6.0% of the genetic diversity. In the present study, this stallion (Dash for Cash) contributed to 11.0% of genetic diversity in the last decade, indicating concentration of genetic origins over the 38 years of competitions, in a small group of ancestors, consequently loss of genetic diversity over the three decades evaluated.

Pure and half-blooded thoroughbred ancestors increased their genetic contribution over the last decades (Table 3). This increase is understood as a search by breeders in using stallions with recent origins from the ancestors of the Thoroughbred. During breeders selection it is common for the pedigree to be the only selection parameter used by breeders. The fact that certain horses present in their ancestors, important Thoroughbred breeders in the formation of the current QH racing animals, being grandchildren, great-grandchildren, cousins or even half-brothers of horses that achieved good performances in races, is a reason enough to be selected as breeders. Another possible reason is the search for lighter (less muscular) and elegant QH horses (greater limb compliance) becoming animals closer to the racial standard of Thoroughbred horses, due to the genetic increase of Thoroughbred ancestors. At the same time, breeders kept pure QH animals. It is possible in this year of 2020, to see in the racetracks pure animals of the QH breed, horses with a morphology of racial standard closer to Thoroughbred and distant from the pure QH of the other two strains (stock and halter / pleasure horses).

Although several studies have demonstrated the loss of genetic diversity in the QH breed (Tunnell et al., 1983; Petersen et al., 2014; Faria et al., 2018a; Marchiori et al., 2019), including the present study, the large number of animals of the QH breed

(more than 2 million registered in the United States plus half a million in Brazil and the remaining around the world) will permit recovery of its genetic diversity.

The pedigree chart (Supplemental Figure S1) generated from the most influential ancestors (Table 3) ends in the last decade and involves more than 39 breeding animals, demonstrating the relatedness between animals. Thus, the American Quarter Horse Association, ABQM and other associations should conduct joint studies designed to identify different founders using desirable selection criteria for racing.

5. Conclusion

Breeders have little control over GIs, with a consequent small genetic gain in athlete horses. The loss of genetic diversity and concentration in a reduced number of breeding animals indicate failures in the breeding and conservation programs of the racing line. The large number of animals on the North American continent whose founders differ from those of the Brazilian line suggests a possible solution for genetic diversity. Mating designs in conjunction with the selection of breeding animals should be promoted and goals should be defined for racing line of the QH breed.

Acknowledgments

The authors wish to thank the Jockey Clube de Sorocaba (JCS, Brazil) and Associação Brasileira de Criadores do Cavalo Quarto de Milha (ABQM) for providing the data used in this study. We also thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes, Brazil) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil). The authors thank the reviewers of Livestock Science for their suggestions.

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil (Finance Code 001). PhD fellowship granted (PROEX- Program of Academic Excellence) (88882.180634/2018-01) to Ricardo

Faria and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil)

productivity grant to Prof. J. Augusto II.

AUTHOR STATEMENT

Ricardo Faria: Writing-Original draft preparation, Methodology, Software, Validation, Investigation, Writing-Reviewing and Editing. **Antônio Vicente:** Validation, Investigation, Writing-Reviewing and Editing. **Alejandra Toro:** Visualization, Investigation. **Augusto II Silva:** Term, Conceptualization, Methodology, Validation, Investigation, Supervision.

6. References

- ABQM, 2019. Associação Brasileira de criadores do cavalo Quarto de Milha [WWW Document]. URL <https://www.abqm.com.br/pt/conteudos/quarto-de-milha/quarto-de-milha-no-brasil> (acessado 3.3.20).
- AQHA, 2015. AQHA Annual Report - 2014 [WWW Document]. Horse Stat. URL [https://www.aqha.com/~media/Files/About/Annual Report/2014/Horse Statistics.ashx](https://www.aqha.com/~media/Files/About/Annual%20Report/2014/Horse%20Statistics.ashx) (acessado 12.10.20).
- Baena, M.M., Gervásio, I.C., Rocha, R. de F.B., Procópio, A.M., de Moura, R.S., Meirelles, S.L.C., 2020. Population structure and genetic diversity of Mangalarga Marchador horses. *Livest. Sci.* <https://doi.org/10.1016/j.livsci.2020.104109>
- Boichard, D., Maignel, L., Verrier, É., 1997. The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.* 29, 5–23. <https://doi.org/10.1186/1297-9686-29-1-5>
- Buttram, S.T., Willham, R.L., Wilson, D.E., Heird, J.C., 1988. Genetics of Racing Performance in the American Quarter Horse: I. Description of the Data. *J. Anim. Sci.* 66, 2791. <https://doi.org/10.2527/jas1988.66112791x>
- Caballero, A., Toro, M., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res.* 75, 331–343. <https://doi.org/10.1017/S0016672399004449>
- Caballero, A., Toro, M. a., 2002. Analysis of genetic diversity for the management of

- conserved subdivided populations. *Conserv. Genet.* 3, 289–299.
<https://doi.org/10.1023/A:1019956205473>
- Cervantes, I., Goyache, F., Molina, A., Valera, M., Gutiérrez, J.P., 2008. Application of individual increase in inbreeding to estimate realized effective sizes from real pedigrees. *J. Anim. Breed. Genet.* 125, 301–310. <https://doi.org/10.1111/j.1439-0388.2008.00755.x>
- Corrêa, M.J.M., Mota, M.D.S., 2007. Genetic evaluation of performance traits in Brazilian Quarter Horse. *J. Appl. Genet.* 48, 145–151.
<https://doi.org/10.1007/BF03194672>
- Duru, S., 2017. Pedigree analysis of the Turkish Arab horse population: Structure, inbreeding and genetic variability. *Animal* 11, 1449–1456.
<https://doi.org/10.1017/S175173111700009X>
- Ekiz, B., Kocak, O., 2005. Phenotypic and genetic parameter estimates for racing traits of Arabian horses in Turkey. *J. Anim. Breed. Genet.* 122, 349–356.
<https://doi.org/10.1111/j.1439-0388.2005.00544.x>
- FAO, 1998. Secondary guidelines for development of national farm animal genetic resources management plans: management of small populations at risk., Food and Agricultural Organization. Roma, Italy.
- Faria, R.A.S., Maiorano, A.M., Bernardes, P.A., Pereira, G.L., Silva, M.G.B., Curi, R.A., Silva, J.A.I.V., 2018a. Assessment of pedigree information in the Quarter Horse: Population, breeding and genetic diversity. *Livest. Sci.* 214, 135–141.
<https://doi.org/10.1016/j.livsci.2018.06.001>
- Faria, R.A.S., Silva, J.A.I.V., 2019. Phenotypic and genetic diversity of the Quarter Horse racing. FCAV – UNESP, Campus de Jaboticabal.
<https://doi.org/dx.doi.org/10.13140/RG.2.2.20682.72649>
- Faria, R.A.S., Vicente, A.P.A., dos Santos, R.I.D.G., Maiorano, A.M., Curi, R.A., Chardulo, L.A.L., Vasconcelos Silva, J.A., 2018b. Genetic Diversity of Lusitano Horse in Brazil Using Pedigree Information. *J. Equine Vet. Sci.* 69, 149–158.
<https://doi.org/10.1016/j.jevs.2018.07.009>
- Fleury, J.J., Neto, J.B.F.C., Alvarenga, M.A., 1989. Results from an embryo transfer programme with Mangalarga mares in Brazil. *Equine Vet. J.* 21, 73–74.
<https://doi.org/10.1111/j.2042-3306.1989.tb04680.x>

- Frankham, R., 1996. Introduction to quantitative genetics (4th edn). Trends Genet. 12, 280. [https://doi.org/10.1016/0168-9525\(96\)81458-2](https://doi.org/10.1016/0168-9525(96)81458-2)
- Garbe, J.R., Da, Y., 2008. Pedigraph: A Software Tool for the Graphing and Analysis of Large Complex Pedigree. User manual Version 2.4. Department of Animal Science, University of Minnesota.
- Giontella, A., Pieramati, C., Silvestrelli, M., Sarti, F.M., 2019. Analysis of founders and performance test effects on an autochthonous horse population through pedigree analysis: Structure, genetic variability and inbreeding. Animal. <https://doi.org/10.1017/S1751731118001180>
- González-Recio, O., López de Maturana, E., Gutiérrez, J.P., 2007. Inbreeding Depression on Female Fertility and Calving Ease in Spanish Dairy Cattle. J. Dairy Sci. <https://doi.org/10.3168/jds.2007-0203>
- Gutiérrez, J.P., Cervantes, I., Goyache, F., 2009. Improving the estimation of realized effective population sizes in farm animals. J. Anim. Breed. Genet. <https://doi.org/10.1111/j.1439-0388.2009.00810.x>
- Gutiérrez, J.P., Cervantes, I., Molina, A., Valera, M., Goyache, F., Olina, A.M., Alera, M. V., 2008. Individual increase in inbreeding allows estimating effective sizes from pedigrees. Genet. Sel. Evol. 40, 359–378. <https://doi.org/10.1051/gse>
- Gutiérrez, J.P., Goyache, F., 2005. A note on ENDOG: A computer program for analysing pedigree information. J. Anim. Breed. Genet. <https://doi.org/10.1111/j.1439-0388.2005.00512.x>
- IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.
- James, J.W., 1972. Computation of genetic contributions from pedigrees. Theor. Appl. Genet. 42, 272–3. <https://doi.org/10.1007/BF00277555>
- Lacy, R.C., 1989. Analysis of founder representation in pedigrees - founder equivalents and founder genome equivalents. Zoo Biol. 8, 111–123. <https://doi.org/10.1002/zoo.1430080203>
- Maciel, F.C., Bertoli, C.D., Braccini Neto, J., Cobuci, J.A., Paiva, S.R., McManus, C.M., 2014. Population structure and genealogical analysis of the Brazilian Crioula Horse. Anim. Genet. Resour. génétiques Anim. génétiques Anim. <https://doi.org/10.1017/s2078633613000489>

- Maignel, L., Boichard, D., Verrier, E., 1996. Genetic Variability of French Dairy Breeds Estimated From Pedigree Information, in: *Interbull Bull.* p. 48–56.
- Mancin, E., Ablondi, M., Mantovani, R., Pigozzi, G., Sabbioni, A., Sartori, C., 2020. Genetic Variability in the Italian Heavy Draught Horse from Pedigree Data and Genomic Information. *Animals* 10, 1310. <https://doi.org/10.3390/ani10081310>
- Marchiori, C.M., Pereira, G.L., Maiorano, A.M., Rogatto, G.M., Assoni, A.D., Augusto II V. Silva, J., Chardulo, L.A.L., Curi, R.A., 2019. Linkage disequilibrium and population structure characterization in the cutting and racing lines of Quarter Horses bred in Brazil. *Livest. Sci.* 219, 45–51. <https://doi.org/10.1016/j.livsci.2018.11.013>
- Meuwissen, T., Luo, Z., 1992. Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* 24, 305–313. <https://doi.org/10.1186/1297-9686-24-4-305>
- Mota, M.D.S., Abrahão, A.R., Oliveira, H.N., 2005. Genetic and environmental parameters for racing time at different distances in Brazilian Thoroughbreds. *J. Anim. Breed. Genet.* 122, 393–399. <https://doi.org/10.1111/j.1439-0388.2005.00551.x>
- Mota, M.D.S., Oliveira, H.N., Silva, R.G., 1998. Genetic and environmental factors that affect the best time of Thoroughbred horses in Brazil. *J. Anim. Breed. Genet.* 115, 123–129. <https://doi.org/10.1111/j.1439-0388.1998.tb00335.x>
- Nielsen, B.D., Turner, K.K., Ventura, B.A., Woodward, A.D., O'Connor, C.I., 2006. Racing speeds of Quarter Horses, Thoroughbreds and Arabians. *Equine Vet. J.* 38, 128–132. <https://doi.org/10.1111/j.2042-3306.2006.tb05528.x>
- Petersen, J.L., Mickelson, J.R., Cleary, K.D., McCue, M.E., 2014. The american quarter horse: Population structure and relationship to the thoroughbred. *J. Hered.* 105, 148–162. <https://doi.org/10.1093/jhered/est079>
- Pjontek, J., Kadlečík, O., Kasarda, R., Horný, M., 2012. Pedigree analysis in four Slovak endangered horse breeds. *Czech J. Anim. Sci.* 57, 54–64.
- Schubertová, Z., Kadlečík, O., Candrák, J., Kasarda, R., 2014. Pedigree analysis of Thoroughbred horses in Slovakia. *Acta Fytotech. Zootech.* <https://doi.org/10.15414/afz.2014.17.04.122-126>
- Siderits, M., Baumung, R., Fuerst-Waltl, B., 2013. Pedigree analysis in the German

- Paint Horse: Genetic variability and the influence of pedigree quality. *Livest. Sci.* 151, 152–157. <https://doi.org/10.1016/j.livsci.2012.10.018>
- Silva, A.P.A., Curi, R.A., Langlois, B., Silva, J.A. V, 2014. Genetic parameters for earnings in quarter horse. *Genet. Mol. Res.* 13, 5840–5848. <https://doi.org/10.4238/2014.August.1.2>
- Stephens, T.D., Splan, R.K., 2013. Genetic variability of the American Shire Horse via pedigree analysis. *J. Equine Vet. Sci.* <https://doi.org/10.1016/j.jevs.2013.03.029>
- Taveira, R.Z., Mota, M.D.S., Oliveira, H.N., 2004. Population parameters in Brazilian Thoroughbred. *J. Anim. Breed. Genet.* 121, 384–391. <https://doi.org/10.1111/j.1439-0388.2004.00472.x>
- Tunnell, J.A., Sanders, J.O., Williams, J.D., Potter, G.D., 1983. Pedigree analysis of four decades of Quarter Horse breeding. *J. Anim. Sci.* 57, 585–593.
- Vostrá-Vydrová, H., Vostrý, L., Hofmanová, B., Krupa, E., Zavadilová, L., 2016. Pedigree analysis of the endangered Old Kladruber horse population. *Livest. Sci.* 185, 17–23. <https://doi.org/10.1016/j.livsci.2016.01.001>
- Wright, S., 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating Author (s): Sewall Wright Reviewed work (s): Published by : Society for the Study of Evolution Stable URL : <http://www.jstor.org/stable/2406450> . *Evolution* (N. Y). 19, 395–420.

TABLES

Table 1. Summary statistics of the animals, pedigree quality, generation interval, Inbreeding parameters (Inbreeding and Average relatedness coefficients, effective population size and Wright's F statistics) to the total racing population and subpopulations of the animals QH with performance in sprint races in Brazil (racing line)

Item	Racing population			
	Total (Pr)	Subpopulation by decade		
		sP80	sP90	sP00
Animals and pedigree quality (n)				
Total	5,861	1,712	1,604	1,461
Males	2,472	610	651	693
Female	3,389	1,102	953	768
Complete equivalent generation traced	5.6 ± 1.2	4.9 ± 0.9	5.7 ± 0.8	6.3 ± 0.9
Generation interval (years)				
Sire-son/daughter	13.2 ± 7.9	15.2 ± 9.2	12.6 ± 8.8	15.1 ± 6.5
Dam-son/daughter	10.6 ± 7.5	12.3 ± 8.3	9.0 ± 7.9	11.4 ± 5.8
Inbreeding parameters (%)				
All animals				
Inbreeding coefficient, F	0.95	0.49	0.91	1.60
Average relatedness coefficient, AR	1.84	1.12	2.05	2.56
Individual increase in inbreeding, ΔF_i	0.22	0.13	0.22	0.34
Animals with F ≠ 0				
Inbreeding coefficient, F	1.55	1.16	1.29	1.80
Average relatedness coefficient, AR	2.35	1.68	2.28	2.71
Individual increase in inbreeding, ΔF_i	0.36	0.26	0.31	0.38
Effective population size, N_e (n)				
N_e based on ΔF_i	215	369	226	144
N_e per generation	68	117	62	41
Wright's F (%)				
F_{IT}	-0.01	-0.01	-0.01	-0.01
F_{IS}	-0.24	-0.08	-0.10	-0.23
F_{ST}	0.23	0.06	0.09	0.22

Pr, total racing population; **sP80**, subpopulation 1980-1989; **sP90**, subpopulation 1990-1999; **sP00**, subpopulation 2000-2009; **F**, inbreeding coefficient; **AR**, average relatedness coefficient; **ΔF_i** , individual increase in inbreeding; **F_{IT}** , individual inbreeding coefficient in relation to the total population; **F_{IS}** , inbreeding coefficient of the individual in relation to its subpopulation; **F_{ST}** , average inbreeding of the subpopulation in relation to the total population.

Table 2. Summary statistics of the founders, probability of gene origin, genetic drift, ancestors, and their genetic contributions to the total racing population and subpopulations of the animals QH with performance in sprint races in Brazil (racing line)

Item	Racing population			
	Total (Pr)	Subpopulation by decade		
		sP80	sP90	sP00
Number of founders, N_f	2,269	1,888	1,451	1,100
Probability of gene origin (n)				
Effective number of founders, f_e	192	209	144	111
Effective number of ancestors, f_a	61	88	44	26
Effective number of founder genome equivalents, f_{ge}	34	55	25	15
Genetic drift (n)				
f_e/N_f	0.08	0.11	0.10	0.10
f_e/f_a	3.15	2.38	3.27	4.27
f_{ge}/f_e	0.18	0.26	0.17	0.14
Ancestors (n)				
100% of genetic diversity	1,587	945	673	462
50% of genetic diversity	32	50	19	9
Marginal genetic contributions of the most influential ancestors (%)				
10 ancestors	32.8	26.8	38.0	52.2
100 ancestors	67.0	61.5	76.2	86.5
500 ancestors	88.6	90.3	97.9	a

Pr, total racing population; **sP80**, subpopulation 1980-1989; **sP90**, subpopulation 1990-1999; **sP00**, subpopulation 2000-2009; **N_f** , effective number of founders; **f_e** , effective number of ancestors; **f_a** , effective number of founder genome equivalents; **f_{ge}** , effective number of founders; **a**, fewer than 500 animals explain the total genetic diversity of the subpopulation (sP00).

Table 3. Summary statistics of the name, sex, year and country of birth, breed, number of offspring, number of equivalent generations, and marginal genetic contributions of the 10 most influential ancestors of the total racing population and of the subpopulations of the animals QH with performance in sprint races in Brazil (racing line)

Animal	Sex	Birth		Breed	Number of offspring	EG	Marginal genetic contribution							
		Year	Country				a	Pr (%)	a	sP80 (%)	a	sP90 (%)	a	sP00 (%)
Dash for Cash	♂	1973	USA	½QH	57	6.1	1	7.04	-	b	2	7.66	1	11.10
Easy Jet	♂	1967	USA	½QH	90	5.0	2	6.28	10	1.21	1	8.88	2	9.20
Beduino	♂	1968	MEX	THO	22	3.5	3	3.75	-	b	3	3.63	3	6.56
Three Bars	♂	1940	USA	THO	44	3.9	4	3.22	1	6.33	5	3.16	10	2.45
Signed to Fly	♂	1988	USA	QH	355	5.3	5	2.92	-	b	4	3.51	4	6.28
Top Deck	♂	1945	USA	THO	18	3.9	6	2.26	2	4.01	6	2.48	-	b
First Prize Rose	♀	1976	USA	QH	1	5.1	7	2.04	-	b	-	b	6	4.25
Streakin Six	♂	1977	USA	QH	32	5.2	8	1.86	-	b	7	2.45	7	3.03
Lady's Moon	♂	1972	USA	QH	145	5.4	9	1.85	3	3.31	8	2.20	-	b
Blazen Bryan	♂	1993	CAN	QH	164	5.1	10	1.61	-	b	-	b	5	4.79
Apollo VM	♂	1988	BRA	QH	190	6.1	-	b	-	b	-	b	8	2.90
Little Smoothie	♀	1973	USA	QH	3	5.0	-	b	-	b	-	b	9	2.54
All	-	-	-	33.33c	93d	5.0	-	32.83e	-	14.86e	-	33.97e	-	53.10e
						d		15.89f		10.95f		17.54f		19.16f

EG, number of complete equivalent generation traced; **a**, order of cumulative marginal genetic contribution; **Pr**, total racing population; **sP80**, subpopulation 1980-1989; **sP90**, subpopulation 1990-1999; **sP00**, subpopulation 2000-2009; **b**, does not appear in the group of the 10 most influential ancestors; ♂, male; ♀, female; **USA**, United States; **MEX**, Mexico; **CAN**, Canada; **BRA**, Brazil; ½ **QH**, ancestor with one Thoroughbred parent and one Quarter Horse parent; **THO**, Thoroughbred ancestor; **QH**, already recognized Quarter Horse ancestor; **c**, percentage of Thoroughbred animals; **d**, average of the 12 ancestors; **e**, sum of the marginal genetic contribution of most influential ancestors in Pr, sP80, sP90 and sP00; **f**, sum of the marginal genetic contribution of ancestors it blood is THO or ½ QH in Pr 5 ancestors (3 THO and 2 ½ QH), in Sp80 3 ancestors (2 THO and 1 ½ QH), in Sp90 5 ancestors (3 THO and 2 ½ QH) in Sp00 4 ancestors (2 THO and 2 ½ QH).

Table 4. Degree of relatedness (%) among the most influential ancestors of the total racing population and of the most recent subpopulation 2000-2009 of the animals QH with performance in sprint races in Brazil (racing line)

Ancestor	A	B	C	D	E	F	G	H	I	J	K	L
Dash for Cash (A)	-	4.99	0.00a	13.48	2.81	6.25	0.81	2.08	1.65	0.45	26.25	7.25
Easy Jet (B)	-	-	0.00a	31.25	26.56	12.50	13.28	25.90	6.64	0.98	27.04	10.40
Beduino (C)	-	-	-	0.00a	0.00a	0.00	0.00a	0.00a	0.00a	25.00	0.00a	0.00a
Three Bars (D)	-	-	-	-	7.81	0.00a	3.13	8.20	9.38	3.13	11.18	13.28
Signed to Fly (E)	-	-	-	-	-	9.38	3.71	6.67	2.05	0.24	7.35	3.38
Top Deck (F)	-	-	-	-	-	-	6.25	3.13	6.25	0.00a	4.69	12.50
First Prize Rose (G)	-	-	-	-	-	-	-	3.34	3.34	0.10	3.53	3.54
Streakin Six (H)	-	-	-	-	-	-	-	-	1.70	0.26	13.25	2.65
Lady's Moon (I)	-	-	-	-	-	-	-	-	-	0.29	2.07	4.37
Blazen Bryan (J)	-	-	-	-	-	-	-	-	-	-	0.36	0.44
Apollo VM (K)	-	-	-	-	-	-	-	-	-	-	-	4.41
Little Smoothie (L)	-	-	-	-	-	-	-	-	-	-	-	-

a, average relatedness coefficient equal to 0.0, however, there are distant common ancestors among the animals.

FIGURE

Supplemental Figure S1. Relationship pedigree chart (partial) generated for the most influential ancestors of the total racing population and of the most recent subpopulation 2000-2009 of the animals QH with performance in sprint races in Brazil (racing line).

Journal Pre-proof



Departamento de Melhoramento e Nutrição Animal – FMVZ – UNESP

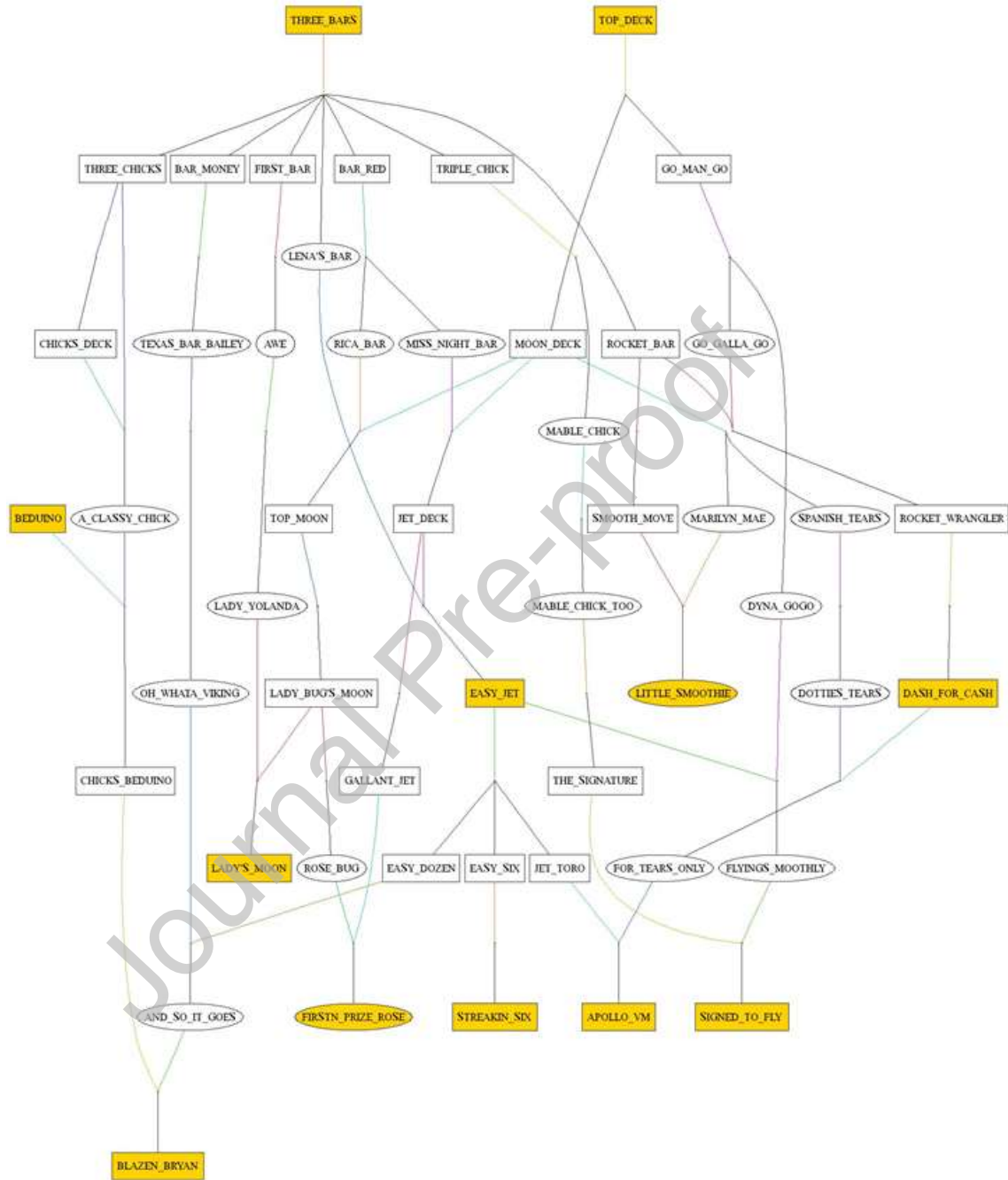
Distrito de Rubião Jr., s/n – Botucatu/SP – 18618-000

☎ - fax: +55-14-3880-0000

✉ augusto.ii@unesp.br

🌐 www.fmvz.unesp.br





Departamento de Melhoramento e Nutrição Animal – FMVZ – UNESP

Distrito de Rubião Jr., s/n – Botucatu/SP – 18618-000

☎ - fax: +55-14-3880-0000

✉ augusto.ii@unesp.br

🌐 www.fmvz.unesp.br

