

Available online at www.sciencedirect.com

Livestock Science 113 (2008) 24–33

**LIVESTOCK
SCIENCE**www.elsevier.com/locate/livsci

Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis

I. Cervantes ^{a,*}, A. Molina ^a, F. Goyache ^b, J.P. Gutiérrez ^c, M. Valera ^d^a Department of Genetics, University of Córdoba, Ctra. Madrid-Cádiz, km 396a, 14071 Córdoba, Spain^b SERIDA-Somió, C/Camino de los Claveles 604, E-33203 Gijón, Spain^c Department of Animal Production, Faculty of Veterinary, UCM, Avda. Puerta del Hierro s/n, 28040 Madrid, Spain^d Department of Agro-Forestal Sciences, EUITA, University of Sevilla, Ctra. Utrera km 1, 41013 Sevilla, Spain

Received 4 September 2006; received in revised form 26 January 2007; accepted 14 February 2007

Abstract

The studbook of the Spanish Arab Horse was analysed in order to assess the available genetic variability in the population via genealogical analyses and to ascertain the influence of different countries in the formation of the population. The Spanish Arab horse studbook could be traced back to the late 19th century with a total of 18,880 animals. A 15.7% (2965) of the registered individuals were Arab horses imported from foreign populations. The total number of founders was 1626: from these, 1495 individuals were identified as imported from foreign countries. The sum of the genetic contributions of the founders to the population revealed that those imported from the 'Middle East', Poland, the United Kingdom and Egypt had a higher importance in the formation of the Spanish Arab Horse, contributing, respectively, 36.4%, 18.7%, 11.9% and 9.1% of the genetic variability. The effective number of founders was 38.6 and the effective number of ancestors was 19 thus characterising the abusive use of a little number of individuals for reproduction. This was due to a breeding policy that included mating between relatives, but also due to the limited availability of reproductive individuals after the Spanish Civil War. The average inbreeding of the population was 7.0% (9.8% for the individuals born during 1995–2004). Individuals with very high levels of inbreeding ($\geq 12.5\%$) were 17.7% in the whole population and 26.7% of the individuals born in the last decade. Positive genealogical F_{IS} near 2% characterised a breeding policy including matings between relatives. No clear within-population differentiation (inferred via genealogical F_{ST}) could be assessed between sire lines ($F_{ST}=0.01$) and no differentiation between geographical areas existed. Overall, the Spanish Arab Horse can be described as a population without genetic subdivision and with a mating policy and historical constraints that reduced the available genetic variability regardless of the relatively large size of its founder population. The obtained information will be useful to implement a selection program in the Spanish Arab Horse.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Spanish Arab horse; Genetic variability; Pedigree analysis; Population history; Within-population differentiation

1. Introduction

The Arab breed is one of the most influential horse breeds in the World (Bowling and Ruvinsky, 2000a). It is distributed worldwide and has been involved in the formation of many other horse breeds, such as the

* Corresponding author. Tel.: +34 957218735; fax: +34 957218707.
E-mail address: v82cenai@gmail.com (I. Cervantes).

Thoroughbred (Bowling and Ruvinsky, 2000b) or the Lipizzan (Zechner et al., 2002). At present, the Arab horse breed has the third census size in Spain, totalling 8.1% of the horses in our country and when the Arab and the Arab crossbreed horse breeds (Hispano-Arab and Anglo-Arab) are considered it moves up to the second place among horse populations in Spain after the Andalusian (Spanish Purebred) horse (MAPyA, 2003). The Spanish Arab studbook was founded in 1847 by the Spanish Ministry of War using, mainly, individuals imported from the Middle East and from various European countries (Maxwell, 1995). During the first third of the 20th century a number of private studs were formed in Southern and Central Spain from individuals purchased in the Military Stud of Jerez de la Frontera and foreign countries such as Poland, the United Kingdom, France, Egypt and the area of the Arabian Peninsula (Fenaux, 1995; Maxwell, 1995). Since the 70's a significant number of Spanish Arab Horse studs have been founded in other Spanish regions such as Catalonia, Balearic Islands and the Cantabrian provinces. The Spanish Arab studbook remains open to include Arab horses from foreign countries that are not always registered with full genealogies in the Spanish Arab Studbook. As what happened in other Spanish horse breeds (Valera et al., 2005) the interest in the Spanish Arab Horse has increased during the past decades with the number of registrations in the studbook. Currently, the Spanish Arab Horse Breeder Association (AECCA) is developing a selection scheme with the aim of improving conformation traits and horse performance, mainly, in endurance rides.

The assessment of the within-population genetic variability, population structure and gene flow is necessary during the implementation of selection programs to establish an appropriate management of the genetic stock in order to enlarge genetic basis for selection. In this respect genealogical tools are important to describe these subjects (Moureaux et al., 1996; Głażewska and Jezierski, 2004). Moreover, demographical analyses may also help to understand important circumstances affecting the genetic history of the population (Zechner et al., 2002; Valera et al., 2005).

The aim of this work is to analyse the pedigree information of the Spanish Arab Horse included in its studbook in order to quantify the available genetic variability in the population in terms of inbreeding levels and founder contributions to the present population. The study will also focus on analysing the influence of animals from foreign countries in the Spanish Arab Horse and to assess a possible population structure. The implications on the genetic history of the

Spanish Arab Horse population over a century of breeding will also be discussed.

2. Materials and methods

Information registered in the studbook of the Spanish Arab Horse from its foundation in 1847 to December 2004 has been analysed. Data included 18,880 (8893 males and 9987 females) registered animals. Current and older studbooks of major Spanish Arab studs were primary sources of information for the identification of the origin of the individuals imported from foreign countries (FC). Important secondary sources were the books of Fenaux (1995), Maxwell (1995) and Steen (2006). Altogether, FC individuals included those animals actually imported, their ancestors included in the Spanish Arab Horse studbook regardless of location in the Spanish territory and those individuals born in Spain but registered as imported 'in uterus'. Those horses which were known to have been bred in Spain before their registration in the studbook and those individuals born in Spain from registered individuals regardless of the country of origin of their parents were classified as Spanish native individuals (SN). Following Fenaux (1995), those horses assigned to the old 'Duque de Veragua' stud, from which genealogical information was lost during the Spanish Civil War, were also considered as SN individuals. Following Moureaux et al. (1996) ancestors with both parents unknown in the available database (managed and provided by the Spanish Ministry of Defence) were considered founders. If one parent is known, the unknown parent was considered a founder (Lacy, 1989). Notice that, using this criterion, some FC individuals can be considered founders for the present study regardless if their parents are registered in foreign Arab horse studbooks.

Traditionally, few familiar lines were identified in the Spanish Arab Horse according to the historical importance of a given stallion in the formation of the population (Fenaux, 1995): *Ursus*, *Seanderich*, *Razada* and *Bagdad*. These sires have been traced forward in the pedigree in order to identify 'sire lines' defined as unbroken descent through male animals only from an ancestor to a descendant (Cunningham et al., 2001; Royo et al., 2006).

The analyses listed below were carried out for the whole population. However some informative parameters were also computed for a group of animals formed by the animals born between 1995 and 2004, in order to approach the last generation. Throughout the text this group of individuals will be named as 'reference population'.

2.1. Demographic parameters

The pedigree completeness level was assessed by computing for the whole pedigree as the proportion of ancestors known per generation for each offspring (MacCluer et al., 1983). Also, the number of equivalent generations (g_e) in the pedigree was computed as the sum of $(1/2)^n$ where n is the number of generations separating the individual to each known ancestor (Boichard et al., 1997).

2.2. Probability of gene origin

The genetic representation of the founder population in the Spanish Arab Horse was assessed using the following parameters:

- Effective number of founders (f_e). This parameter is the reciprocal of the probability that two genes drawn at random in the studied population originate from the same founder (James, 1972) and is computed from the genetic contribution of founders to the descendant gene pool of the population (Lacy, 1989).
- Effective number of ancestors (f_a). To compute this parameter, the ancestors explaining a percentage of population higher than their parents were identified, and only their marginal contribution, that one not explained by other ancestors previously chosen, was considered. This parameter complements the information offered by the effective number of founders accounting for the losses of genetic variability produced for the unbalanced use of reproductive individuals producing bottlenecks. It could be defined as the minimum number of ancestors, not necessarily founders, explaining the complete genetic diversity of a population (Boichard et al., 1997). The genetic contributions of the ancestors have been averaged by decade of birth of the ancestors in order to identify populations' bottlenecks.

2.3. Inbreeding and average relatedness coefficient

The level of heterozygosity of the population was assessed using the following parameters:

- Individual inbreeding coefficient (F) defined as the probability that an individual has two genes identical by descent (Malécot, 1948).
- Average relatedness coefficient (AR) of each individual defined as the probability that an allele randomly chosen from the whole population belongs

to a given animal (Goyache et al., 2003; Gutiérrez et al., 2003). The AR coefficient for each individual in the pedigree is computed as the average of the coefficients in the row correspondent to the individual in the numerator relationship matrix **A** (Gutiérrez et al., 2003). Notice that both alleles from the individual are included in the population. AR is twice the mean coancestry between a given animal and all animals in the population including itself. Thus the AR accounts simultaneously for the coancestry and inbreeding coefficient. Notice that the AR coefficient of a founder is its genetic contribution to the entire pedigree.

In order to ascertain the genetic contribution (Lacy, 1989) over the time of the historically sire lines, we traced forward the genetic representation of *Ursus*, *Seanderich*, *Razada* and *Bagdad* in all the individuals of the population and averaged them by decade of birth of their descendants.

2.4. Genetic structure

Genetic structure of the Spanish Arab Horse population was assessed using F -statistics (F_{IT} , F_{ST} , F_{IS} ; Wright, 1978). F_{IT} is the inbreeding coefficient of the individual relative to the whole population, F_{ST} is the average inbreeding of the subpopulation relative to the whole population and F_{IS} is the inbreeding coefficient of the individual relative to its own subpopulation (Falconer and Mackay, 1996). These parameters were computed following Caballero and Toro (2000, 2002) as $F_{IT} = \frac{\bar{F} - \bar{f}}{1 - \bar{f}}$, $F_{ST} = \frac{\bar{f} - \bar{f}}{1 - \bar{f}}$, and $F_{IS} = \frac{\bar{F} - \bar{f}}{1 - \bar{f}}$, where \bar{f} and \bar{F} are, respectively, the mean coancestry and the inbreeding coefficient for the entire population, and \bar{f} the average coancestry for the subpopulation so that $(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$. This analysis was carried out considering two different levels of population subdivision: the sire line level and the geographical area level. At the sire line level the analysis included the traditional *Ursus*, *Seanderich*, *Razada*, and *Bagdad* lines besides the 'Unsigned' stallions as subpopulations. At the geographical area level the analysis was carried out considering as independent populations: a) the Spanish Mediterranean Levant and Balearic Islands; b) Northern Spain; c) Andalusia; d) Centre of Spain; and e) the Military Stud which was treated as an independent population because of its breeding importance.

Most parameters have been computed using the program ENDOG v 3.0 (Gutiérrez and Goyache, 2005) and other programs developed by the authors.

3. Results

3.1. Demographic analyses

Table 1 gives the frequency of individuals registered in the Spanish Arab studbook per year of birth (grouped in decades) of the animals. Thirty three percent of the individuals were registered during the last decade (1995–2004), illustrating a demographic evolution caused by an increasing interest of the breeders jointly with the improvement of recording activity. The number of registrations decreased during the decade 1935–44 due to the consequences of the Spanish Civil War (1936–1939). Among the registered individuals, the proportions of males and females used for reproduction were 12.6% and 27.9%, respectively.

Considering the progeny size, most stallions (71.2%) had an offspring size between 1 and 5. However, the use of the males for reproduction is not balanced and a little number of sires (27; 1.1% of the stallions) had more than 70 descendants.

The total number of FC (Foreign Countries) individuals registered in the Spanish Arab studbook was 2965 (15.7 % of the total). The contribution of the FC individuals to the studbook is high; a 47.9% of the SN (Spanish Native) individuals registered during the period 1905–1934 had FC parents. During the period 1955–2004 this proportion was of 16.4% (Table 1), in spite the fact that most of FC animals (74.5%) were imported between 1945 and 1994. The total number of founders was 1626, 1495 of them being FC founders.

The founders of Spanish origin totalled 131 individuals. The country of origin of 333 FC founder individuals could not be identified with certainty. The other founder FC individuals (1162) were imported from a total of 29 countries. The percentage of founder individuals registered in the studbook according to their country of origin is shown in Fig. 1 (see Plot A). The United Kingdom was the country with the highest number of FC founders individuals (12.5%) followed by Poland (10.1%). The other countries that significantly contributed with founder FC individuals to the Spanish Arab studbook were: USA, Egypt, ‘Middle East’ (which is the area including the Arabian Peninsula, and the present Jordan, Syria and Iraq) and France, supplying, respectively, 9.3%, 8.5%, 6.9% and 5% of the FC founders individuals. The SN founders were 8.1% of the total.

The quality of the available pedigree information is illustrated in Fig. 2. Completeness level for the whole pedigree was of 92.0%, 86.6% and 80.8% for, respectively, the first, second and third known generations. These levels drop to values below 40% after the eighth generation. When only the animals born in the reference population are considered the completeness of the pedigree information was higher than 90% until the sixth parental generation. In any case no more than 8 full generations (both parents known) could be traced for any individual in the pedigree. The situation above is also reflected by the number of equivalent generations; this parameter was 7.9 for the individuals born and registered in the last decade whilst for the whole pedigree reached values of 5.7.

Table 1
Description of the Spanish Arab Horse population by decade of birth of the individuals registered in the studbook

| Decade | Total registrations | Number individuals with foreign parents (%) | Genetic contribution ^a | AR ^b | F ^c | F ₅ ^d |
|-----------|---------------------|---------------------------------------------|-----------------------------------|-----------------|----------------|-----------------------------|
| <1895 | 86 | 86 (100) | 0.3 | 0.2 | 0 | 0 |
| 1895–1904 | 172 | 25 (14.5) | 12.5 | 0.4 | 0 | 0 |
| 1905–1914 | 158 | 105 (66.5) | 7.6 | 1.9 | 0.1 | 0.1 |
| 1915–1924 | 434 | 262 (60.4) | 5.1 | 3.0 | 1.3 | 1.3 |
| 1925–1934 | 791 | 296 (37.4) | 38.2 | 4.4 | 2.2 | 2.1 |
| 1935–1944 | 615 | 65 (10.6) | 22.1 | 4.8 | 2.2 | 2.1 |
| 1945–1954 | 1014 | 84 (8.3) | 3.3 | 4.7 | 2.5 | 2.4 |
| 1955–1964 | 1095 | 159 (14.5) | 2.9 | 5.3 | 3.2 | 2.9 |
| 1965–1974 | 1313 | 370 (28.2) | 3.3 | 6.2 | 3.7 | 3.1 |
| 1975–1984 | 2766 | 635 (23.0) | 3.2 | 9.2 | 6.4 | 4.9 |
| 1985–1994 | 4196 | 595 (14.2) | 1.1 | 11.2 | 9.3 | 6.1 |
| 1995–2004 | 6240 | 805 (12.9) | 0.4 | 11.6 | 9.8 | 5.2 |

^a Average genetic contribution of the Boichard et al.'s (1997) ancestors (%).

^b Mean average relatedness values (%).

^c Average inbreeding computed using the total pedigree (%).

^d Average inbreeding computed using only the last 5 generations (%).

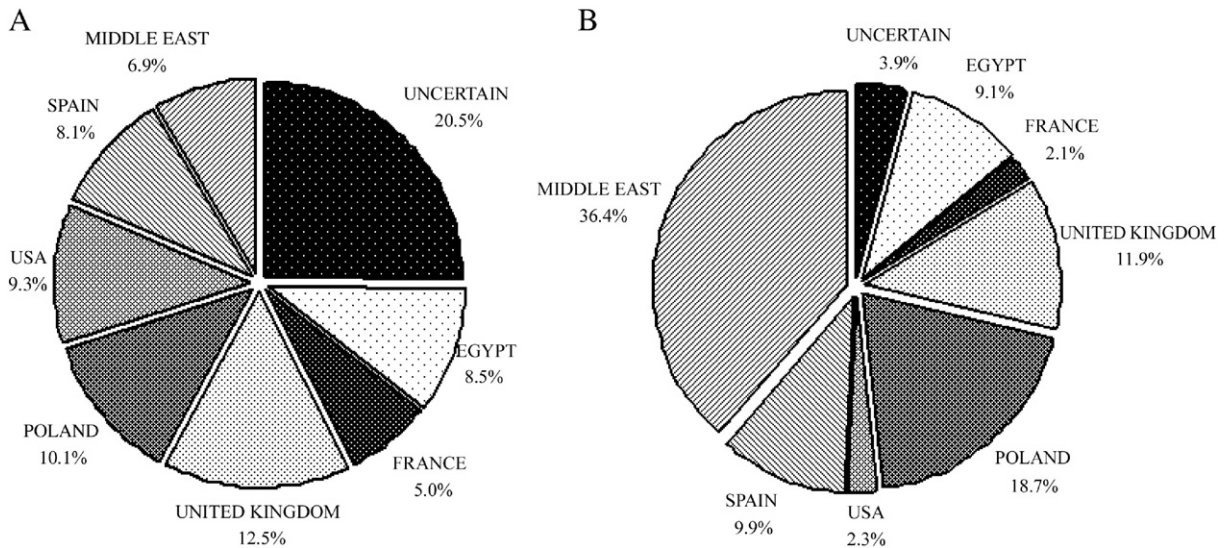


Fig. 1. Contributions (in percentage) of different countries (or areas) of origin to the Spanish Arab studbook. Plot A shows the percentage of founder individuals registered in the studbook whilst Plot B gives the sum of the contributions to the genetic variability of the population of these founder individuals. The founder individuals with no clear origin have been classified as ‘Uncertain’.

3.2. Probability of gene origin

Parameters characterising the genetic variability of the population in terms of gene origin for the whole and the reference populations are given in Table 2. A total of 1626 (1385 with both parents unknown and 241 with one parent unknown) individuals were identified as founders. The effective number of founders for the whole population (38.6) was particularly small accounting for a 2.4% of the founders.

The genetic contribution of the FC founders was summed up by country of origin to quantify the importance of each country as a source of genes in the formation of the Spanish Arab Horse (Fig. 1, see Plot B). FC founders with its country of origin identified with certainty contributed to the genetic variability of the

population in a large extent (86.0%) whilst the contribution of the SN founders was of 9.9%. Geographical areas contributing significantly to the genetic variability of the population were ‘Middle East’ (36.4%), Poland (18.7%), the United Kingdom (11.9%) and Egypt (9.1%). However, other countries such as USA or France that contributed to the Spanish Arab Horse with a significant number of founder individuals summed up contributions lower than 3% (2.3% and 2.1%, respectively), due to the very small contribution of these founder animals.

The number of ancestors (founders or not) that explained 100% of genetic variability of the whole population was 994, although only 7 individuals were necessary to explain 50% (see Table 2). For the reference population these figures were respectively 417

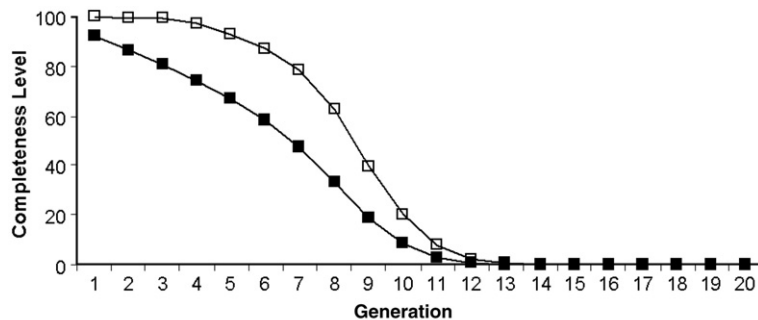


Fig. 2. Average percentage of ancestors known per generation, (generation 1 = parents, generation 2 = grandparents, etc.) in Spanish Arab Horse. Information is provided for the whole population (in black squares) and for the last decade (open squares).

Table 2

Parameters characterising the probability of gene origin in the Spanish Arab Horse population for the whole pedigree and for the reference population (animals born in the last decade)

| | Whole population | Reference population |
|-----------------------------------------------------------------------|------------------|----------------------|
| Total number of animals in the population | 18,880 | 6240 |
| Animals with both parents unknown | 1385 | 762 |
| Animals with one parent unknown | 241 | 98 |
| Number of founders | 1626 | 860 |
| Effective number of founders | 38.6 | 39.5 |
| Effective number of ancestors | 19 | 13 |
| Number of ancestors explaining 100% genetic variability of population | 994 | 417 |
| Number of ancestors explaining 50% genetic variability of population | 7 | 5 |

and 5. The effective numbers of ancestors were 19 and 13 respectively for the whole and the reference populations. Table 1 gives the sum of the contributions of the identified Spanish Arab ancestors according to their decade of birth. Major contributions were given by ancestors born during the decades 1895–1904 (12.5%), 1925–1934 (38.2%) and 1935–1944 (22.1%) whilst those born in the other periods have a sum of the contributions lower than 8%. This characterises the dramatic scenario suffered by the population during and after the Spanish Civil War period in which a few number of individuals were massively used for reproduction.

Table 3 shows the contribution of the 15 main ancestors (identified according to Boichard et al., 1997) identified for the whole and the reference populations. Ancestors with the higher contribution to the Spanish Arab Horse were *Gandhy* (13.8% and 16.8%, respectively, for the whole and the reference populations) and *Congo* (12.7% and 16.3%, respectively, for the whole and the reference populations).

Fig. 3 shows the contribution of the historically important stallions (sire lines) to the population per decade of birth of their descendants. The stallion *Ursus* was the most influential with an increasing importance over decades to finally reach values of 14.0% in the animals born during the last decade. The stallions *Seanderich* and *Razada* contributed moderately to the population (7.2% and 5.0%, respectively, during the last decade). Finally, *Bagdad* has shown to have a low genetic influence on the registered individuals (over 2% from 1945 to present).

3.3. Inbreeding and average relatedness coefficients

The average values of F and AR for the whole analysed pedigree were, respectively, 7.0% and 9.1%.

The variation of average F and AR values by decades of birth or the animals is given in Table 1. The average F values were higher than 1% from the 10's being higher than 6% from 1975. This parameter reached average values of 9.8% in the reference population. During the last decade the number of inbred individuals is important. Table 4 gives the frequency of Spanish Arab Horse individuals per degree of inbreeding. Individuals with very high levels of inbreeding ($F \geq 12.5\%$) were 17.7% in the whole population and more than a quarter (26.7%) of the registered individuals born in the last decade. The variation of the mean AR values was similar to that of F with maximum values in the last decade of 11.6%. In order to distinguish between close and remote inbreeding the variation of the inbreeding by decades of birth

Table 3

Description of the 15 main ancestors (identified according to Boichard et al., 1997) contributing the most to the Spanish Arab Horse population as a whole and in the last decade

| | Sex | Origin country | Year of birth | Genetic contribution (%) |
|-----------------------------|-----|----------------|---------------|--------------------------|
| <i>Whole population</i> | | | | |
| 1. Gandhy | M | Spain | 1931 | 13.8 |
| 2. Congo | M | Spain | 1941 | 12.7 |
| 3. Nana-sahib | M | Spain | 1934 | 7.2 |
| 4. Wan-dick | M | Poland | 1901 | 6.7 |
| 5. Barquillo | M | Spain | 1938 | 4.1 |
| 6. Yaima | F | Spain | 1935 | 3.4 |
| 7. Sara 1930 | F | Spain | 1930 | 3.4 |
| 8. Veracruz | F | Spain | 1934 | 2.8 |
| 9. Razada | M | England | 1925 | 2.0 |
| 10. Korosko | M | Egypt | 1907 | 1.9 |
| 11. Seanderich | M | Middle East | 1902 | 1.8 |
| 12. Veranda | F | Spain | 1933 | 1.7 |
| 13. Reyna | F | England | 1925 | 1.6 |
| 14. Sabat-el-heir | M | Egypt | 1906 | 1.5 |
| 15. Ursus | M | Poland | 1908 | 1.4 |
| <i>Reference population</i> | | | | |
| 1. Gandhy | M | Spain | 1931 | 16.8 |
| 2. Congo | M | Spain | 1941 | 16.3 |
| 3. Nana-sahib | M | Spain | 1934 | 7.8 |
| 4. Barquillo | M | Spain | 1938 | 6.7 |
| 5. Yaima | F | Spain | 1935 | 5.3 |
| 6. Sara 1930 | F | Spain | 1930 | 4.3 |
| 7. Veracruz | F | Spain | 1934 | 3.9 |
| 8. Ocalina | F | Spain | 1951 | 2.7 |
| 9. Abisinia 1935 | F | Spain | 1935 | 2.3 |
| 10. Wan-dick | M | Poland | 1901 | 2.0 |
| 11. Veranda | F | Spain | 1933 | 1.9 |
| 12. Nazeer | M | Egypt | 1934 | 1.5 |
| 13. Saboya | F | Spain | 1931 | 1.4 |
| 14. Ghazel | F | Middle East | 1924 | 1.2 |
| 15. Kashmir | M | Spain | 1934 | 1.1 |

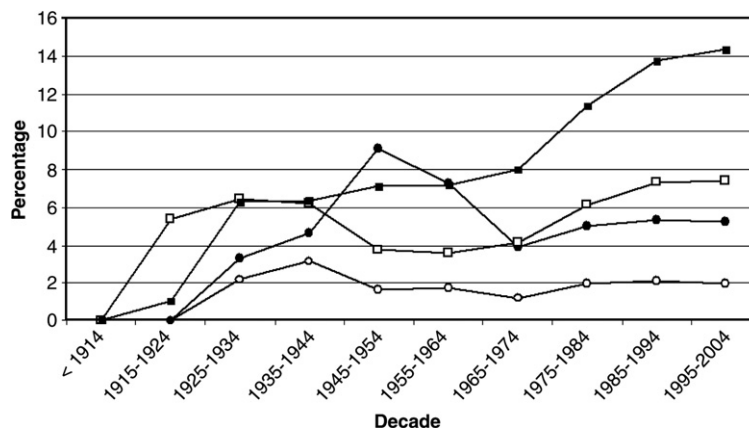


Fig. 3. Influence of the sire lines *Ursus* (black squares), *Seanderich*, (open squares) *Razada* (black circles) and *Bagdad* (open circles) in the Spanish Arab Horse by decade.

computed using only data from the last 5 (F_5) generations was also given (Table 1). Patterns of variation of F and F_5 were consistent till the decade 1955–1964 when they reached values around 3%. After this decade, the average F_5 of the registered individuals increased more slowly than F to reach maximum values of 6.1% in 1985–1994 whereas in the last decade decreased to 5.2%. However the maximum value of F was 9.8% for the same period.

3.4. Within-population subdivision

Within-population genetic differentiation was assessed using F -statistics computed from genealogies. This analysis was carried out at two different population levels considering: a) the sire line level; and b) the geographical area level (see Material and methods section). At the sire line level the statistics F_{ST} reached values of $F_{ST}=0.010$, $F_{IS}=0.016$ and $F_{IT}=0.025$. However, no differentiation was assessed by geographical area with values of the corresponding F -statistics of $F_{ST}=0.001$, $F_{IS}=0.019$ and $F_{IT}=0.025$.

Table 4

Absolute and relative (in percentage) frequency of individuals registered in Spanish Arab Horse by rate of inbreeding (F)

| Rate of inbreeding | Whole population | | Decade 1995–2004 | |
|--------------------------|------------------|------|------------------|------|
| | N | % | N | % |
| Non inbred ($F=0$) | 5007 | 26.5 | 550 | 8.9 |
| $0 < F < 6.25\%$ | 3709 | 19.6 | 956 | 15.4 |
| $6.25\% \leq F < 12.5\%$ | 6817 | 36.1 | 3048 | 49.0 |
| $F \geq 12.5\%$ | 3347 | 17.7 | 1660 | 26.7 |

Data are shown by the whole data set and for the last decade.

4. Discussion

4.1. Demographic analyses

The Arab horse is widely spread in the whole world founding particular populations in many countries. The Spanish Arab Horse was mainly formed using individuals from foreign countries and the studbook remains open today for FC Arab individuals.

The sum of the genetic contributions of the founder individuals to the population according to the country of origin makes possible to ascertain major sources of genes in the history of the Spanish Arab Horse. Here we show that the major source of genes of the Spanish Arab Horse is ‘Middle East’ coinciding with the historical major breeding area of the breed. Moreover, the importance of the Arab horse populations of Poland, the United Kingdom, and Egypt in the formation of the Spanish Arab has been also confirmed. Major genetic contributions by country of origin assessed here roughly coincide with the type of animals bred in the Spanish Arab studs (Fenaux, 1995): traditionally, they select individuals of ‘Egyptian type’ to improve type traits whilst individuals of ‘Polish type’ are preferred to improve movements.

The Spanish Arab Horse is a population with historical pedigrees that can be traced back to the mid 19th century (Maxwell, 1995). In this respect, the whole Spanish Arab Horse pedigree has lower level of completeness than other comparable breeds due to the constant inclusion of FC individuals in the studbook. The Spanish Arab Horse, as a whole, had a completeness level higher than 74% when up to four generations are traced back. The Andalusian horse breed, with comparable pedigree depth, had completeness about

90% until this generation (Valera et al., 2005). In Lipizzan horse a level of completeness of 90% was observed until the tenth generation (Zechner et al., 2002). This relative lack of pedigree information has also been highlighted by the number of equivalent generations which is, for the whole Spanish Arab Horse population (5.7), lower than the values reported in Lipizzan (15.2; Zechner et al., 2002) or Andalusian (8.3; Valera et al., 2005). However, the quality of the available genealogical information has substantially increased during the last decades in which the interest of the breeders and the number of registrations in the studbook have increased significantly. In this respect, in our reference population the level of completeness is higher than 90% until the sixth generation, similar to the value for the French Arab, over 90% in the fourth generation (Moureaux et al., 1996).

4.2. Genetic variability

The assessment of the genetic variability of a population is a necessary step before the implementation of an improvement program. This will be determined by the size of the base population but also by the mating policy and population structure that will allow different degrees of genetic contributions from founder to the present population. The number of founders in the reference population of the Spanish Arab Horse (860) is lower to that previously reported for the French Arab (962 founders; Moureaux et al., 1996). However, this kind of comparisons must be interpreted with caution since the ratio number of founders over total number of animals depends on the time from foundation of the Studbook and the quality of pedigree data recording. The total number of founders (1626) of the Spanish Arab Horse population (that are 8.6% of the registered individuals) is similar to that recently reported for the Andalusian horse breed (1465 founders) though in the later breed the founders account for roughly a 2% of the total population (Valera et al., 2005). This is due to the main characteristic of the Spanish Arab studbook that remains open to other Arab horse population from its foundation to the present.

However, parameters characterising the genetic contributions from founders to the population do not reach higher values thus reflecting an unbalanced use of some individuals for reproduction which affects the whole genetic stock. The effective number of founders, which measures the overall founder representation in a managed population accounting for the loss of genetic variability from unequal founder contributions (Lacy, 1989, 1995) was 39.5 for the reference population in

Spanish Arab Horse (4.6% of total founders), which is lower than that of 135 (14.0% of total founders) obtained in French Arab Horse (Moureaux et al., 1996) and other French horse breeds such as Trotteur français (70; 2.8% of total founders) or Selle français (333; 3.7% of the founders) as reported by Moureaux et al. (1996). However, the analyses by Moureaux et al. (1996) differ from those presented here in levels of completeness, total number of founders and number of generations used for the study, and, in consequence, comparisons are not straightforward. In any case, the effective number of founders for the whole population (38.6, 2.4% of total number of founders) in the Spanish Arab Horse is also lower than others recently reported in horse breeds with comparable pedigree depths such as Lipizzan (48.2; 10.55% of total founders; Zechner et al., 2002) or Andalusian (39.6; 2.7%; Valera et al., 2005) but higher than that reported for Thoroughbred (28; 17.8%; Cunningham et al., 2001). The effective number of ancestors (which is here 19 for the whole population, roughly half of the effective number of founders) that accounts for bottlenecks in the population confirms the abusive use of a small number of individuals for reproduction in the Spanish Arab Horse. Its value in Spanish Arab Horse is lower than that of 26.2 obtained for the whole pedigree in Lipizzan horses (Zechner et al., 2002) and similar than that of 16.5 obtained for the whole pedigree in Andalusian horse (Valera et al., 2005) that have more complete and deeper genealogies registered in the studbook.

Information given in Tables 2 and 3 is useful to characterise the breeding policy of the population. The individual contribution to the genetic variability of the population from the individuals identified as ancestors (Boichard et al., 1997) is considered; the genetic contribution of the major two Spanish Arab ancestors to the whole population (13.8% and 12.7%) is quite similar to that reported in other horse breeds with unbalanced mating policies such as Andalusian horse (the major two ancestors contributing 15.8% and 12.6% respectively). In fact, only 7 Spanish Arab ancestors are needed to explain 50% of the genetic variability of the population whilst in the Andalusian horse breed and in the Lipizzan population 6 and 10 ancestors, are respectively needed (Valera et al., 2005; Zechner et al., 2002). The overall information reported here characterises the Spanish Arab horse as a population that, regardless the high number of founders registered in the studbook, has followed a mating policy allowing for the very unbalanced use of some individuals. This can have a historical reason that can arise from the genetic bottleneck suffered by the population due to the

Spanish Civil War (Table 1). Genetic contributions of the ancestors born during the period of the War and the previous decade summed up approximately 50%. Most major ancestors described in Table 3 were born during these decades and, in fact, we can consider that the Spanish Arab Horse was genetically ‘re-founded’ during the late 30’s and the 40’s.

4.3. Inbreeding

The average F value computed for the whole pedigree of Spanish Arab Horse population is lower than that computed in other Spanish populations as Andalusian horse (8.5%; Valera et al., 2005). This can be explained by the special scenario described in our Arab horse in which the studbook is still open for the registration of Arab individuals. The pedigree of the new registered individuals is not always completely included in the Spanish Arab Studbook, producing subsequently animals considered as founders. However, we had expected lower levels of inbreeding than those assessed in this scenario in which the rate between the number of founder individuals and the total population size exceed that of breeds with comparable pedigree depth. The average F computed using the reference population (9.8%) is higher than those of 3.1% and 5.3% reported, respectively, by Moureaux et al. (1996) and Głazewska and Jeziński (2004) in French and Polish Arab horses, for the individuals born during the more recent period in their databases. This fact and the proportion of individuals with levels of inbreeding higher than 12.5% (see Table 4) point out that breeders have planned most matings using related individuals. Recently, Głazewska and Gralak (2006) have shown in Polish Arab horse that breeders have carried out a balancing selection favouring the more heterozygous mares (at the protein polymorphism level) for reproduction. This can cause discrepancies between real gene pool in a population and that estimated on the basis of pedigrees (Głazewska and Gralak, 2006; Royo et al., 2006).

Moreover, after 1975 average F is higher than half mean AR thus pointing out that matings are mainly carried out between relatives. This is also reflected by the parameter F_{IS} . The F_{IS} statistic computed from genealogies is equivalent to the parameter α (Caballero and Toro, 2000; Toro et al., 2000), which means the departure from random mating as a deviation from Hardy–Weinberg proportions. Positive genealogical F_{IS} values mean that the average F within a subpopulation exceeds between-individuals coancestry (Gutiérrez et al., 2005) and occurs if mating between relatives is not avoided (Toro et al., 2000). F_{IS} values computed at

the sire line level and at geographical area level reached positive values near 2% (0.016 and 0.019, respectively). However, this breeding policy might have been changed these past few years. The trends of the inbreeding computed using only the last 5 generations (F_5) showed that during the last decade the average F_5 was roughly 1% lower than that computed in the ante last decade. F depends to a large extent on the quality of pedigree information especially in the initial generations registered. F_5 still increases in our Arab horse population up to 1994 accounting for the mating policy described above. However in the last decade a major effort seems to have been made to correct the possible effect of such high levels of inbreeding in the population.

The wide use of particular individuals in the population is also reflected in the lack of within-population genetic differentiation assessed using genealogical F_{ST} s. At the sire line level a little differentiation over 1% was obtained due to the predominance of some sire genealogies in particular studs. However, stallions assigned to a given paternal line share genes (inherited by their remaining ancestors) with the other sire lines. So, traditional assignment of a stallion to a sire line in the Spanish Arab Horse (Maxwell, 1995) does not have clear genetic support. Moreover, there is no geographical structure in our Arab horse. The expansion of the breed from Andalusia to other parts of Spain has not been made by using particular sources of genes; neither from FC nor SN individuals. The breeding significance of the Military Stud of Jerez de la Frontera has probably caused this lack of geographic differentiation. In fact, the Military stud continues to be the most significant source of genes in the Spanish Arab Horse. This governmental stud sets up public mounting services throughout the Spanish territory using its stallions those that are used in most medium–small studs.

5. Conclusions

Here we carried out the first analysis of the pedigree information registered in the Spanish Arab Horse studbook to assess the genetic variability and the population structure. The information obtained will be useful to implement a selection program in the Spanish Arab Horse. The high number of founders registered in the studbook would suggest the existence of a large genetic basis for a stallion’s selection program. However, historical constraints, such as the consequences of the Spanish Civil War, and the breeding policy carried out by the breeders have reduced the available genetic variability as reflected in small values computed for parameters such as the effective number of founders and

the effective number of ancestors. In the light of our results a selection program in the Spanish Arab Horse would need an appropriate management of the genetic stock. The implementation of mating policies promoting the unbalanced use of those individuals with lower AR values for reproduction, as proposed previously for other Spanish horse breeds (Valera et al., 2005) and the exchanges between Arab populations from different countries can be useful to maintain the genetic variability of the population and also to enlarge genetic basis for selection.

Acknowledgements

The authors thank the support given by the ‘Subdirección General de Medios de Producción Ganaderos’ of the Spanish Ministry of Agriculture. This work has been made within a collaboration agreement between the Spanish Arab horse breeders association (AECCA) and the University of Sevilla. We also thank the ‘Fondo de Explotación de los Servicios de Cría Caballar y Remonta’ of the Spanish Ministry of Defence for its collaboration.

References

- Boichard, D., Maignel, L., Verrier, E., 1997. The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.* 29, 5–23.
- Bowling, A.T., Ruvinsky, A., 2000a. The genetic resources and their conservation. *The Genetic of the Horse*. CAB International, UK.
- Bowling, A.T., Ruvinsky, A., 2000b. Genetic aspects of domestication. *The Genetic of the Horse*. CAB International, UK.
- Caballero, A., Toro, M.A., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res. Camb.* 75, 331–343.
- Caballero, A., Toro, M.A., 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* 3, 289–299.
- Cunningham, E.P., Dooley, J.J., Splan, R.K., Bradley, D.G., 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Anim. Genet.* 32, 360–364.
- Falconer, D.S., Mackay, F.C., 1996. *Introduction to Quantitative Genetics*. Longman Group Ltd., England.
- Fenaux, K., 1995. *El Pura Sangre Árabe*. Ediciones el Caballo, Barcelona, España.
- Głazewska, I., Jezierski, T., 2004. Pedigree analysis of Polish Arabian Horses based on founder contributions. *Livest. Prod. Sci.* 90, 293–298.
- Głazewska, I., Gralak, B., 2006. Balancing selection in Polish Arabian Horses. *Livest. Sci.* 105, 272–276.
- Goyache, F., Gutiérrez, J.P., Fernández, I., Gómez, J., Álvarez, I., Diez, J., Royo, L.J., 2003. Using pedigree information to monitor genetic variability of endangered population: the Xalda sheep breed of Asturias as an example. *J. Anim. Breed. Genet.* 120, 95–103.
- Gutiérrez, J.P., Goyache, F., 2005. A note on ENDOG: a computer program for analysing pedigree information. *J. Anim. Breed. Genet.* 122, 172–176.
- Gutiérrez, J.P., Altarriba, J., Díaz, C., Quintanilla, R., Cañón, J., Piedrafito, J., 2003. Genetic analysis of eight Spanish beef cattle breeds. *Genet. Sel. Evol.* 35, 43–64.
- Gutiérrez, J.P., Marmí, J., Goyache, F., Jordana, J., 2005. Pedigree information reveals moderate to high levels of inbreeding and a population genetic structure in the Catalanian donkey breed. *J. Anim. Breed. Genet.* 122, 378–386.
- James, J., 1972. Computation of genetic contributions from pedigrees. *Theor. Appl. Genet.* 42, 272–273.
- Lacy, R.C., 1989. Analysis of founder representation in pedigrees: founder equivalent and founder genome equivalents. *Zool. Biol.* 8, 111–123.
- Lacy, R.C., 1995. Clarification of the genetic terms and their use in the management of captive populations. *Zool. Biol.* 14, 565–578.
- MacCluer, J., Boyce, B., Buke, L., Weitzkamp, D., Pfenning, A., Parsons, C., 1983. Inbreeding and pedigree structure in Standardbred horses. *J. Heredity* 74, 394–399.
- Malécot, G., 1948. *Les Mathématiques de l’Hérédité*. Masson et Cie, Paris.
- Maxwell, J., 1995. *Spanish Arabian Horse Families*. Alexander Heriot and Co. Ltd, England.
- Moureaux, S., Verrier, E., Ricard, A., Meriaux, J.C., 1996. Genetic variability within French race and riding horse breeds from genealogical data and blood marker polymorphisms. *Genet. Sel. Evol.* 28, 83–102.
- MAPYA, 2003. Estudio y caracterización genética del sector equino en España. Ministerio de Agricultura Pesca y Alimentación. (<http://www.mapa.es/app/Equino/Informacion/Infsector.aspx?lng=Des>). España.
- Royo, L.J., Álvarez, I., Gutiérrez, J.P., Fernández, I., Goyache, F., 2006. Genetic variability in the endangered Asturcón pony assessed using genealogical and molecular information. *Livest. Sci.* 107, 162–169. doi:10.1016/j.livsci.2006.09.010.
- Steen, A.K., 2006. *A Glorious Task. A Tale of the Breed*. Escandón Impresores. Seville, Spain.
- Toro, M.A., Rodríguez, J., Silió, M., Rodríguez, C., 2000. Genealogical analysis of a close herd Black Hairless Iberian pigs. *Conserv. Biol.* 14, 1843–1851.
- Valera, M., Molina, A., Gutiérrez, J.P., Gómez, J., Goyache, F., 2005. Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. *Livest. Prod. Sci.* 95, 57–66.
- Wright, S., 1978. *Evolution and the genetics of populations. Variability Within and among Natural Populations*, vol. 4. University of Chicago Press, Chicago, USA.
- Zechner, P., Sölkner, J., Bodo, I., Druml, T., Baumung, R., Achmann, R., Marti, E., Habe, F., Brem, G., 2002. analysis Diversity and population structure in the Lipizzan horse bred based on pedigree information. *Livest. Prod. Sci.* 77, 137–146.