A genetic analysis of the Italian Salernitano horse

A. Criscione¹, V. Moltisanti¹, L. Chies², D. Marletta¹† and S. Bordonaro¹

¹Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A) Università degli Studi di Catania, via Valdisavoia 5, 95123 Catania, Italy; ²Dipartimento di Scienze delle Produzioni Agrarie, Università Mediterranea di Reggio Calabria, Località Feo di Vito, 89122 Reggio Calabria, Italy

(Received 20 January 2015; Accepted 8 May 2015; First published online 6 July 2015)

Salernitano (SAL) is an ancient Italian horse breed developed over the course of the ages together with Napoletano and, during the 20th century, by crossing with Thoroughbred horse lines. Excellent in hurdle jumping, this breed is currently facing a concrete risk of extinction due to the lack of appropriate management strategies. This research is the first SAL genetic characterization that aims to set up the basic knowledge for a conservation plan. A representative sample of 61 SALs was analyzed by means of a set of 16 microsatellites markers (short tandem repeats (STRs)). The sequence of hypervariable D-loop mtDNA region was also performed on a subset of 24 mares in order to study the maternal diversity and obtain a complete picture of the internal genetic variation. All the molecular data were analyzed together with those obtained from three Sicilian horse breeds investigated in a previous research (Siciliano, Sanfratellano and Sicilian Oriental Purebred). STRs markers revealed a moderate level of genetic diversity in SAL (alleles/locus 5.1, He 0.67) and confirmed the hunch of genetic erosion. Autosomal variability highlighted a very light deficit of homozygotes (FIS = −0.067). Experimental D-loop sequences were compared by multiple alignments with those retrieved from biological databases and revealed two unreported haplotypes. The phylogenetic network, which was built on mtDNA sequences, included various cosmopolitan and European horses and showed SAL haplotypes distributed among different mtDNA lineages.

Keywords: Salernitano horse, genetic diversity, microsatellites, D-loop mt-DNA, breed conservation

Implications

Salernitano (SAL) is a saddle horse that used to be a model of excellence in equestrian sport. For instance, between 1956 and 1972, SAL horses were winners at the Olympic Games. Today, this ancient Italian breed is severely endangered and counts <100 horses. This research is focused on the analysis of the genetic characterization of the currently existing SAL breed. The aim is to provide useful information to plan suitable strategies for SAL management and conservation. The maintenance of this portion of horse biodiversity may help preserve an important resource of equestrian sector as well as positively exploit the role of horse breeding in local areas.

Introduction

The Salernitano (SAL) is one of the oldest Italian horse breeds developed in the rural plains of the Campania region in Southern Italy, an area already considered by the Romans exceptionally suited for horse breeding (Hendricks, 1995). The original strain was crossed with Oriental and Andalusian horses (Chiari, 1901) and developed together with Napoletano breed that greatly arose in reputation in 1532, when the Federico Grisone’s riding academy was established in Naples (Difflouch, 1923). Starting from 1780s, SAL began to be selectively bred at the Persano Stud in Piana del Sele (province of Salerno, Campania). During the 20th century, SAL was crossed with Thoroughbred in order to enhance its beauty of shapes and sport abilities. Between the years 1956 and 1972, excellent SAL horses (Merano, Posillipo and Fiorello) took the lead at International Show Jumping Championships and Olympic Games. The current existing SAL is a mesodolicomorphic saddle horse that still excels in hurdle jumping. Today, this breed is facing a concrete risk of extinction: in 2015, <100 horses were listed in the Book of equine population with limited diffusion (MIPAAF Ministerial decree n.552/2009) managed by Italian Breeder Association (Associazione Italiana Allevatori, AIA). It is noteworthy that till January 2015 AIA used to list SAL together with Persano (188 horses in total), using the identification (Persano-Salernitano) for these two different breeds (AIA, 2015). A genetic analysis of this endangered breed is a prerequisite for a plan to safeguard its germplasm. The aim of this paper is the investigation of the SAL’s biodiversity and its evaluation in the context of Southern Italy breeds as well as among European and cosmopolitan horse breeds. The use

1 E-mail: d.marletta@unict.it
of different molecular approaches, implementing nuclear (short tandem repeats (STRs)) and mtDNA (D-loop) markers, can provide broad information of genetic diversity to address suitable strategies for management and conservation of the ancient SAL breed.

**Material and methods**

**Sampling and DNA extraction**

Peripheral blood samples (10 ml in K3-EDTA tubes) were collected from 61 SAL horses in one experimental herd (stud of ARSSA in Cassano Allo Ionio, Cosenza, Italy). DNA was extracted using the commercial Illustra blood genomic Prep Mini Spin kits (GE Healthcare, Little Chaifont, UK).

A subset of 24 mares was chosen trying to avoid tightly related animals in order to investigate the D-loop mtDNA sequence diversity.

**Markers’ analysis**

Genetic characterization was carried out for a set of 16 STRs (HTG6, HTG10, VH20, HTG7, HTG4, AHT5, AHT4, HMS3, HMS54, HMS2, ASB2, HMS8, HTG15, HMS1, HMS5), which were amplified and analyzed by a 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The individual genotypes were assessed by the use of GeneScan® Analysis ver. 3.1.2 and Genotyper® ver. 2.5 (Applied Biosystems) software. Mitochondrial control region (D-loop), which were amplified and sequenced using the BigDye Terminator v1.1 Kit, on a 3130 Genetic Analyzer.

A subset of 24 mares was chosen trying to avoid tightly related animals in order to investigate the D-loop mtDNA sequence diversity.

**Statistical analysis**

The STR markers obtained from STRs analysis were analyzed together with those from 191 Sicilian horses (80 Siciliano (SIC), 61 Sanfratellano (SAN) and 50 Sicilian Oriental Purebred (SOP)) investigated in previous research by Guastella et al. (2011) through the same panel of microsatellite markers. The matching of allele sizes was performed by using the same reference samples. The main parameters of microsatellite variability were calculated using the GENALEX ver. 6.1 (Peakall and Smouse, 2006) software. Polymorphic information content (PIC) was calculated according to Botstein et al. (1980). The allelic richness (AR), the distribution of genetic variability assessed through F-statistics (FIT, FIS and FST) and the pairwise FST distances were estimated through the FSTAT ver.2.9.3 software (Goudet, 2001). Nei’s (D_A) genetic distance (Nei et al., 1983) that assumes genetic differentiation due to mutation and genetic drift, was calculated using Population software ver. 1.2.30 (Langella et al., 1999). The model-based approach proposed by Falush et al. (2003) in the STRUCTURE 2.2 software was used to evaluate the genomic clustering of the sample. The admixture model was implemented to infer the population structure using no prior information (100 000 burn-ins, 100 000 iterations). The number of tested clusters (K) ranged from 1 to 7 (10 runs for each K). The software CLUMPP ver. 1.1.2 (Jakobsson and Rosenberg, 2007) was implemented to find the optimal alignment of the 10 replicate cluster analyses of the same K. The mean membership matrix across replicates was plotted with the program DISTRUCT ver.1.1 (Rosenberg, 2004).

Using the NETWORK ver.4.6.1.2. software, the mitochondrial DNA diversity of SAL, SAN, SOP, SIC and 173 sequences (11 cosmopolitan and European breeds) retrieved from National Center for Biotechnology Information GenBank were also investigated (Accession numbers are given in Supplementary Table S1) through the use of the maximum parsimony network (Polzin and Daneschmand, 2003), built using the median-joining method (Bandelt et al., 1999).

**Results**

**Microsatellites (STRs)**

The STR markers showed a good informative power (PIC = 0.613 in SAL and 0.715 in the set of four Italian breeds) with PIC per locus that ranged from 0.543 to 0.820 in the whole sample of 252 horses. In SAL, ASB2, HTG15 and HMSS loci showed low PIC values, 0.388, 0.428 and 0.473, respectively (see Supplementary Table S2).

SAL population showed a moderate autosomal polymorphism with a low average number of alleles (5.1) and expected heterozygosity (0.670). Ar was lower in SAL (5.0) than in SOP, SAN and SIC (5.9, 6.1 and 6.8, respectively) following the same trend of He; no private allele was found in SAL. Observed heterozygosity (0.714) was higher than the expected. A moderate degree of genetic differentiation among the four Italian breeds (FST index 8.5%) was observed. The measures of heterozygote deficit within populations (FIS) and in the global meta-population (FIT) were −0.013 and 0.074, respectively.

According to Nei’s distance (D_A) and the pairwise FST values, SAL was the most differentiated breed in the set. According to both these parameters of differentiation, the highest values were detected between SAL and SOP.
(0.290 and 0.167, respectively), while SAL breed was more similar to SIC (0.131 and 0.085) than the other breeds (Table 1).

Clustering analysis (Figure 1) inferred $K = 4$ as the most probable genomic grouping of the four breeds; each cluster was mainly represented by one breed with percentages of membership which ranged from 68.5% in SIC to 94.7% in SOP, with intermediate values for SAN and SAL (79.7% and 94.6%, respectively). When the admixture analysis inferred $K = 2$ clusters, as expected because of its different origin, only SAL grouped separately; at $K = 3$, SOP differentiated from the Sicilian meta-sample, probably due to the high internal homogeneity and the different genetic origin. Increasing the number of inferred clusters ($K = 5$ to 6) the within breed level of admixture increased in both SAN and SIC; in contrast SAL and SOP genomes still clearly referred to distinct clusters.

### Mitochondrial sequences

In SAL mares, eight haplotypes and 18 polymorphic sites (7.3% on average) were identified (Table 2). All mtDNA D-loop sequences determined in this study (Hap 1-8) were deposited in the GenBank with the accession numbers (KR013114 to KR013119, KJ815028 to KJ815029).

Maternal variability in SAL was higher than in SOP, but lower than in SIC and SAN (Guastella et al., 2011). Among the Southern Italy breeds, SAL counted six exclusive haplotypes out of a total of 26: in this context SAL shared the two most frequent mtDNA sequences with SOP and SAN (Hap1) and with SIC (Hap2), respectively.

The Hap1 showed the highest frequency in SAL (29%) as well as in the whole set of four Italian breeds, as it was the only one found in SOP (20 horses) and is also present in SIC (25%).

The analysis of the molecular diversity appointed SAL less variable than the Sicilian breeds, except for SOP (Guastella et al., 2011) in terms of haplotype diversity ($h = 0.84 \pm 0.04$) and polymorphic sites ($p_s = 18$). Mean number of pairwise differences ($\pi$) and nucleotide diversity ($\pi_n$) values were also estimated in SAL ($\pi = 7.183$; $\pi_n = 0.030 \pm 0.02$) as well as in SIC ($\pi = 7.127$; $\pi_n = 0.031 \pm 0.02$) and SAN ($\pi = 6.515$; $\pi_n = 0.028 \pm 0.01$).

BLAST search showed that four out of six exclusive haplotypes found in the SAL sample overlapped with the GenBank-derived sequences, whereas two sequences were new in the explored databases.

### Discussion

SAL is an ancient and prestigious horse breed. In the last century, many of these excellent horses were used for the show jumping. In 1950s, two SAL horses, ridden by a legendary jumper, the major Raimondo D’Inzeo, won the World Show Jumping Championships (1956) and the Summer Olympic Games (1960).

Despite its famous and relevant history, this breed is facing a concrete risk of extinction and today counts <100 horses. At the beginning of the 70s the interest toward SAL declined: the genetic selection addressed to sport ability was abandoned and the breed was outclassed by the German, French, Dutch and Belgian horses.

This research, carried out on a broad sample of the SAL population, is focused on a genetic analysis performed by means of different molecular markers. Owing to the current circumstances of the breed risk of extinction, this important local genetic resource of Southern Italy showed a moderate molecular variability. The main parameters of genetic variability obtained from the analysis of STRs autosomal markers are quite low. The values found here are similar to those reported by Matassino et al. (2013) in a different sample of SAL (alleles/locus 6.2; $He = 0.689$) and lower than those observed in Maremmano (Felicetti et al., 2010), Danish
The calculated low number of alleles is probably due to the severe reduction in census. Admixture analysis and $D_A$ genetic distances showed a very clear differentiation of the SAL breed in the context of the local breeds of Southern Italy. The overall FST for the whole set of Italian breeds (8.5%) is comparable to those reported in other European horse population studies, which never showed values below 8% (Glowatzki-Mullis et al., 2005; Marletta et al., 2006; Luis et al., 2007), and higher than in four Basque-Navarrese semiferal native horse breeds (Solis et al., 2005).

$D_A$ distance and pairwise FST indicated SAL closer to SIC than the other Southern Italy breeds; this outcome is probably due to the common relationship SAL and SIC had with Thoroughbred in 20th century.

Although the comparison among D-loop haplotypes does not often provide breed identity in horse species, the analysis of mtDNA haplotypes can offer relevant information about the history and the genetic variation of maternal lines in autochthonous populations (Aberle et al., 2007; Moridi et al., 2012).

In SAL, it was observed a moderate genetic base of maternal lines, compared with other horse breeds (Cothran et al., 2005; Pérez-Gutiérrez et al., 2008). Average nucleotide diversity values appointed SAL as genetically more diverse than other European, such as Lusitano and Sorraia (Lopes et al., 2005), and Asian (Kakoi et al., 2007) endangered breeds. Haplotype diversity ($h$) was lower than those reported in SIC, SAN (Guastella et al., 2011) and in seven Italian breeds recently investigated by Bigi et al. (2014).

Despite its condition of endangered breed, SAL revealed an unexpected high variability in the maternal lineages; this finding is in agreement with those recently reported in two Sardinian autochthonous horses (Morelli et al., 2014). It is therefore considered that the bottleneck has affected to a reduced extent the maternal lines. Within the set of four breeds from Southern Italy (SAL, SAN, SOP and SIC) SAL shares its two most frequent haplotypes (Hap1 with 29% and Hap2 with 25%) with SOP and SAN, respectively. These two haplotypes, were widely represented at the worldwide scale. The Hap1, previously reported as unique haplotype in SOP and present in SAN (Guastella et al., 2011), was found in seven SAL mares; in light of these evidences, this haplotype is the most frequent in the breeds of South Italy. This mtDNA sequence overlaps with GenBank sequences belonging to many widespread breeds such as Arab, Barb, Andalusian, Thoroughbred, Lusitano and in three Italian breeds (Bigi et al., 2014). The presence of this haplotype in SAL horses might confirm the assumption that at least one maternal lineage of SAL originated in very ancient times from Oriental mares.

The Hap2, the second most represented haplotype in SAL, is worldwide shared and was also found in other Italian breeds such as SIC (Guastella et al., 2011), Murgese, Maremmano, Ventasso and Haflinger (Bigi et al., 2014).

In SAL the Hap3 and Hap4 had the frequencies of 17% and 13%, respectively, whereas the other four SAL haplotypes native (Thirstrup et al., 2008) and South European horses (Solis et al., 2005; Marletta et al., 2006).

## Table 2: Nucleotide substitutions in eight D-loop haplotypes (397 bp fragment) and distribution (n) detected in Salernitano horse breed, defined by multiple alignment with the reference sequence (Acc num X79547)

| Nucleotide substitution | n | 15494 | 15495 | 15496 | 15534 | 15538 | 15585 | 15597 | 15601 | 15602 | 15603 | 15617 | 15635 | 15649 | 15650 | 15659 | 15666 | 15670 | 15703 | 15709 | 15720 |
|-------------------------|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| X79547                  |   | T     | T     | A     | C     | A     | G     | T     | T     | G     | A     | G     | T     | C     | T     | T     | T     | T     | T     | T     |
| Hap1                    | 7 | C     | C     | G     | T     | A     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     |
| Hap4                    | 3 | .     | .     | .     | .     | A     | .     | .     | .     | .     | .     | .     | .     | A     | .     | .     | .     | .     | .     | .     |
| Hap5                    | 1 | .     | .     | .     | .     | A     | .     | C     | T     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     |
| Hap8                    | 1 | C     | C     | .     | T     | .     | A     | .     | T     | C     | .     | .     | .     | .     | .     | .     | .     | .     | .     | .     |

1Nucleotide substitution at positions 15703, 15709 and 15720 are added to 1 in the median-joining network.
were found as singletons (4% each). Hap3, 4, 5 and 6 overlap with GenBank sequences and were previously observed in other breeds (Bigi et al., 2014). Conversely two sequences were new, so that had been deposited in the database: Hap7 (GenBank Acc Num KJ815028) is terminal in the network, whereas Hap8 (GenBank Acc Num KJ815029), interestingly, belongs to the branch that connects two groups of the network.

The Hap8 represents an ‘ancestral’ haplotype that is essential for the network construction. The median-joining network (Figure 2) highlighted a sharp dichotomy within the maternal variability in SAL breed with the clear differentiation between two haplogroups: one characterized by the CCG motif (Hap1; nt 15494–15496) and the other by the TCA motif (Hap 2–7). Hap8, which has a CCA motif, connects these two branches of the network and therefore it does not

Figure 2. Median-joining network built on a total of 258 haplotypes belonging to 84 Italian autochthonous horses analyzed in this study plus 174 sequences of European and cosmopolitan horse breeds. Pie charts are proportional to the frequency of each haplotype. Small gray circles indicate missing haplotypes; nodes are connected by numbers indicating the position of mutations. Colors represent the breeds as given in the figure. The numbers from 1 to 8 indicate the haplotypes found in Salernitano (KR013114 to KR013119, KJ815028 to KJ815029). SAL = Salernitano; SAN = Sanfratellano; SIC = Siciliano; SOP = Sicilian Oriental Purebred; ARA = Arab; THO = Thoroughbred; AND = Andalusian; BAR = Barb; SOR = Sorraia; LUS = Lusitano; CHE = Cheju; TUV = Tuva; MON = Mongolian; YUM = Yumman; PRZ = Przewalski and reference sequence X79547.
belong to the group including most of western Mediterranean breeds (Iberian, Sicilian and North African horses) nor to the group that comprises the highest frequencies of ‘Arab’ haplotypes (according to Jansen et al., 2002). It is not easy to go back to when this sequence originated, but it surely has a very ancient origin being in the middle of two main haplogroups representing all the modern and the ancient D-loop horse diversity. However, being the only one found so far seems to have had little adaptive success.

The distribution of the matrilineal genetic variation in SAL seems to confirm that: (1) D-loop genetic diversity generally lacks the biogeographical patterning (Vila et al., 2001); (2) D-loop variability has a very ancient origin that probably predates horse domestication (Lippoldt et al., 2011).

In the framework of the global horse biodiversity, D-loop haplotypes here reported in SAL breed have been compared with the equine mtDNA haplogroup nomenclature proposed by Achilli et al. (2012). SAL haplotypes 1, 2, 3 and 5 belong to the haplogroups L, M, G and N, respectively; and their frequencies only partially agreed with those reported in Italian breeds (Bigi et al., 2014). In addition, taking into account the geographic distribution of the haplogroups described in Europe, Asia and Middle East by Achilli et al. (2012), interestingly haplogroup L showed in SAL a frequency (29%) more similar to those reported in the Middle East breeds (22.40%) than in the European ones (38.06%). Finally, in this context it is noteworthy that the high frequency of the haplogroup M (25%), rarely represented in both Italian and worldwide breeds; and the frequency of the group G (17%) that is very similar to that found in Asia (16.35%). These findings together suggest once again the Eastern origin for this breed, strengthening the historical records.

Starting from the 70s, SAL population has been suffering a severe reduction in population. Moreover, its genetic distinctiveness and the breed identity were threatened by the use, for long time, of a common studbook together with the Persiano horse, although these two breeds are different racial units (Matassino et al., 2013). Finally, just recently, the MIPAAF decreed by DM n. 1598/2015 the formal separation of these two ancient breeds with the respective standards. In this paper, molecular data obtained from nuclear (STRs) and mitochondrial (D-loop) markers proved the clear genetic differentiation of SAL among the Italian autochthonous breeds, in agreement with the preliminary evidences reported by Matassino et al. (2013). Furthermore, the presence of eight shared and original D-loop sequences is noteworthy. SAL is an endangered breed that needs to be maintained and preserved by genetic erosion. The results reported here could be useful to address the selection strategy of SAL horse breed. The D-loop diversity, characterized by peculiar female lineages, represents the genetic base from which to plan an appropriate management of mating. Notwithstanding the low number of registered stallions, the selection of reference sires in planned matings is the prerequisite to recover the most important equestrian traits. The maintenance of SAL horse may help preserve an important genetic resource of equestrian sector as well as positively exploit the role of horse breeding in Southern Italy.

Acknowledgments
The authors are grateful to PhD Antonio Zuccaro for the helpful discussion and Dr Carmela Scollo for collecting blood samples. This research was supported by Regione Sicilia –UO n.25 Ricerca applicata e sperimentazione.

To supplementary material for this article, please visit http://dx.doi.org/10.1017/S1751731115001019

References


