PROOF COVER SHEET

Author(s): Maria Cristina Cozzi, Paolo Valiati, Raffaele Cherchi, Erica Gorla, Raphaelle Teresa Matilde Maria Prinsen, Maria Longeri, Alessandro Bagnato, and Maria Giuseppina Strillacci

Article title: Mitochondrial DNA genetic diversity in six Italian donkey breeds (Equus asinus)

Article no: IMDN_A_1292505

Enclosures: 1) Query sheet 2) Article proofs

Dear Author,

1. Please check these proofs carefully. It is the responsibility of the corresponding author to check these and approve or amend them. A second proof is not normally provided. Taylor & Francis cannot be held responsible for uncorrected errors, even if introduced during the production process. Once your corrections have been added to the article, it will be considered ready for publication.

Please limit changes at this stage to the correction of errors. You should not make trivial changes, improve prose style, add new material, or delete existing material at this stage. You may be charged if your corrections are excessive (we would not expect corrections to exceed 30 changes).

For detailed guidance on how to check your proofs, please paste this address into a new browser window: http://journalauthors.tandf.co.uk/production/checkingproofs.asp

Your PDF proof file has been enabled so that you can comment on the proof directly using Adobe Acrobat. If you wish to do this, please save the file to your hard disk first. For further information on marking corrections using Acrobat, please paste this address into a new browser window: http://journalauthors.tandf.co.uk/production/acrobat.asp

2. Please review the table of contributors below and confirm that the first and last names are structured correctly and that the authors are listed in the correct order of contribution. This check is to ensure that your name will appear correctly online and when the article is indexed.

Sequence	Prefix	Given name(s)	Surname	Suffix
1		Maria Cristina	Cozzi	
2		Paolo	Valiati	
3		Raffaele	Cherchi	
4		Erica	Gorla	
5		Raphaelle Teresa Matilde Maria	Prinsen	
6		Maria	Longeri	
7		Alessandro	Bagnato	
8		Maria Giuseppina	Strillacci	

Queries are marked in the margins of the proofs, and you can also click the hyperlinks below.

General points:

- 1. **Permissions:** You have warranted that you have secured the necessary written permission from the appropriate copyright owner for the reproduction of any text, illustration, or other material in your article. Please see http://journalauthors.tandf. co.uk/permissions/usingThirdPartyMaterial.asp.
- 2. Third-party content: If there is third-party content in your article, please check that the rightsholder details for re-use are shown correctly.
- 3. Affiliation: The corresponding author is responsible for ensuring that address and email details are correct for all the coauthors. Affiliations given in the article should be the affiliation at the time the research was conducted. Please see http:// journalauthors.tandf.co.uk/preparation/writing.asp.
- 4. **Funding:** Was your research for this article funded by a funding agency? If so, please insert `This work was supported by <insert the name of the funding agency in full>', followed by the grant number in square brackets `[grant number xxxx]'.
- 5. Supplemental data and underlying research materials: Do you wish to include the location of the underlying research materials (e.g. data, samples or models) for your article? If so, please insert this sentence before the reference section: 'The underlying research materials for this article can be accessed at <full link>/ description of location [author to complete]'. If your article includes supplemental data, the link will also be provided in this paragraph. See <htp://journalauthors.tandf.co. uk/preparation/multimedia.asp> for further explanation of supplemental data and underlying research materials.
- 6. The **PubMed** (http://www.ncbi.nlm.nih.gov/pubmed) and **CrossRef databases** (www.crossref.org/) have been used to validate the references. Changes resulting from mismatches are tracked in red font.

AUTHOR QUERIES

- Q1: Please provide the missing department name for author's affiliation [b].
- Q2: Please check whether the author names (first name followed by last name) and affiliations are correct as presented in the proofs.
- Q3: The citation Ref. (Matassino et al., 2007) has been changed to Ref. (Matassino et al. 2014) as given in the reference list. Please check.
- Q4: Please provide the volume number for Ref. (Bordonaro et al. 2012) in the reference list entry.
- Q5: Please provide the publisher name for Ref. (Starkey & Starkey, 2000) in the reference list enrtry.
- Q6: Artwork will be produced in colour only in the online version, please ensure that Figures 2 and 5 will make sense in black and white for print version, amending captions and text as necessary.
- Q7: Please resupply figure 3 and 5 in a format suitable for printing.

How to make corrections to your proofs using Adobe Acrobat/Reader

Taylor & Francis offers you a choice of options to help you make corrections to your proofs. Your PDF proof file has been enabled so that you can mark up the proof directly using Adobe Acrobat/Reader. This is the simplest and best way for you to ensure that your corrections will be incorporated. If you wish to do this, please follow these instructions:

1. Save the file to your hard disk.

2. Check which version of Adobe Acrobat/Reader you have on your computer. You can do this by clicking on the Help" tab, and then About".

If Adobe Reader is not installed, you can get the latest version free from http://get.adobe.com/reader/.

3. If you have Adobe Acrobat/Reader 10 or a later version, click on the Comment" link at the right-hand side to view the Comments pane.

4. You can then select any text and mark it up for deletion or replacement, or insert new text as needed. Please note that these will clearly be displayed in the Comments pane and secondary annotation is not needed to draw attention to your corrections. If you need to include new sections of text, it is also possible to add a comment to the proofs. To do this, use the Sticky Note tool in the task bar. Please also see our FAQs here: http://journalauthors.tandf.co.uk/production/index. asp.

5. Make sure that you save the file when you close the document before uploading it to CATS using the Upload File" button on the online correction form. If you have more than one file, please zip them together and then upload the zip file. If you prefer, you can make your corrections using the CATS online correction form.

Troubleshooting

Acrobat help: http://helpx.adobe.com/acrobat.html Reader help: http://helpx.adobe.com/reader.html

Please note that full user guides for earlier versions of these programs are available from the Adobe Help pages by clicking on the link Previous versions" under the Help and tutorials" heading from the relevant link above. Commenting functionality is available from Adobe Reader 8.0 onwards and from Adobe Acrobat 7.0 onwards.

Firefox users: Firefox's inbuilt PDF Viewer is set to the default; please see the following for instructions on how to use this and download the PDF to your hard drive: http://support.mozilla.org/en-US/kb/view-pdf-files-firefox-without-downloading-them#w_using-a-pdf-reader-plugin

RESEARCH ARTICLE

1

2

3

4

5

6

7

8

9

10

11 12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

59

O1

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114 115

116

117

118

Mitochondrial DNA genetic diversity in six Italian donkey breeds (Equus asinus)

Maria Cristina Cozzi^a, Paolo Valiati^a, Raffaele Cherchi^b, Erica Gorla^a, Raphaelle Teresa Matilde Maria Prinsen^a, Maria Longeri^a, Alessandro Bagnato^a and Maria Giuseppina Strillacci^a

^aDepartment of Veterinary Medicine, Università degli Studi di Milano, Milano, Italy; ^bAgenzia per la ricerca in agricoltura–AGRIS Sardegna, 02 Sassari, Italy

ABSTRACT

Donkeys have played an important role in agricultural land practices and in human historical periods of recent past and, still today, are used as a working power in several world areas. The objective of this study was to identify genetic variability in six Italian donkey breeds using mtDNA D-loop. Fifteen haplotypes, grouped in three haplogroups, were identified. The genetic indices were informative and showed a high population genetic variability. The results of AMOVA analyses based on geographic structuring of Italian populations highlighted that the majority of the observed variance is due to differences among samples within breeds. Comparison among Italian haplotypes and mtDNA D-loop sequences belonging to European domestic and Ethiopian donkeys and wild asses, clearly define two clades referred to Nubian lineage. The results can be useful to complement safeguard planes for donkey breeds that are considered to extinction endangered.

ARTICLE HISTORY

Received 30 November 2016 Revised 31 January 2017 Accepted 5 February 2017

KEYWORDS

Italian donkeys; mitochondrial DNA; D-loop; biodiversity; genetic variability

Introduction

The domestication of animals largely have contributed to the improvement of the living conditions of the primitive communities and had fundamental effects on human societies. Particularly, the history and origins of the donkey are of strong interest given that, like the horse, the donkey played an important role in agricultural culture and it is still used as a working power in several world rural areas.

The family of the Equidae, including the horses, donkeys or asses, and zebras, was extensively studied during the past two centuries. According to currently accepted International Union for the Conservation of Nature and Natural Resources IUCN taxonomy (Moehlman 2002), modern equids are represented by the eight extant species of the Equus genus: domestic horse (Equus caballus), Przewalski's horse (Equus przewalskii), kiang (Equus kiang), Asiatic wild ass (Equus hemionus), African wild ass (Equus africanus), mountain zebra (Equus zebra), plains zebra (Equus quagga) and Grevy's zebra (Equus grevyi). The domestic donkey (Equus africanus asinus) therefore is considered a sub-specie of African wild ass and many authors suggested that its domestication occurred in the arid tropical and subtropical territories of North East Africa (Blench 2000; Beja-Pereira et al. 2004). Despite that, other studies keep open the debate on the domestication and evolutionary history of the donkeys. In fact, Rossel et al. (2008) analyzing archeological remains excavated in Egypt, strongly supported the African origin of the domestic donkeys.; Marshall (2007) proposed on the other hand that early northeast African pastoralists domesticated the donkey under

conditions of increased aridity 7000-6500 BP in the Sahara desert.

Beja-Pereira et al. (2004) in a worldwide genetic study on modern donkeys using mitochondrial DNA identified two distinct wild African ass sub-species: the Nubian wild ass (Equus africanus africanus) and the Somali wild ass (Equus africanus somalicus). According to this study, the authors found two main clade: the first (clade 1) included the domestic donkeys related to the Nubian wild ass: the second one (clade 2) involved donkeys descending from the Somali wild ass. The authors described high levels of genetic diversities in both the lineages of Northeast Africa territories and suggested that this area is one of the primary centres of donkey domestication. However, a study by Kimura et al. (2011) reported that the mtDNA sequences obtained from the extant Somali wild ass were classified together with the previously identified Somali wild ass specimens, but failed to show any sequence similarity with domestic donkeys of both clades. The study from Kimura et al. (2011) suggested the existence in Northeast Africa territories of another ancestor of the domestic donkeys of clade 2, belonging to an additional yet unrecognized extinct wild population. In a recent study by Kefena et al. (2014) based on genetic and on linguistic and zooarchaeological evidences (Marshall 2007), confirmed that Ethiopia could be one of the major hotspots of donkey diversity and domestication.

Despite its uniquely maternal origin, nowadays, mtDNA nucleotide sequences are a recognized tool for resolving phylogenetic relationships at different evolutionary level,

57 58 CONTACT Maria Cristina Cozzi 🖾 cristina.cozzi@unimi.it 🗊 Department of Veterinary Medicine, Università degli Studi di Milano, Via Celoria 10, 20133, Milano, Italv

© 2017 Informa UK Limited, trading as Taylor & Francis Group

B Supplemental data for this article can be accessed here.

because of its special properties such as: more copies than 119 nuclear DNA, well-known gene structure, lack of introns, 120 high-mutation rate, absence of recombination events (Xu 121 et al. 1996). As such previous studies on donkey mitochon-122 drial DNA D-loop (control region) sequences well depicted 123 the genetic relationships and origin of worldwide modern 124 donkey breeds (Oakenfull et al. 2000; Ivankovic et al. 2002; 125 126 Aranguren-Mendez et al. 2004; Beja-Pereira et al. 2004; Lopez et al. 2005; Chen et al. 2006; Lei et al. 2007; Han et al. 2014; 127 Kefena et al. 2014; Pérez-Pardal et al. 2014; Cinar Kul et al. 128 2016). 129

In Italy donkeys have been bred for centuries for various 130 131 purposes. Historical and archeological evidence suggest that this livestock specie was present and already exploited by 132 humans from around 2000 BC (Kugler et al. 2007-2008). 133 However, during the last century donkey populations size, 134 135 have suffered a decrease caused by the increased mechaniza-136 tion of agriculture and the depopulation of rural districts that 137 reduced the need for donkeys and mules as draft animals. 138 According to FAO (http://dad.fao.org/), some Italian autoch-139 thonous donkey breeds are extinct. In particular, today in Italy eight donkey breeds classified as critically endangered 140 141 (Asinara, Pantesco, Grigio Siciliano and Romagnolo), or endan-142 gered (Amiata, Sardo, Martina Franca and Ragusano) are still 143 bred. However, genetic studies on the biodiversity Italian 144 donkey breeds are limited, and mainly focused on the donkey 145 genetic variability performed using protein markers and 146 microsatellites (Cristofalo et al. 1994; Cosseddu et al. 2001; 147 Ciampolini et al. 2007; Guastella et al. 2007; Matassino et al. 148 2014; Bordonaro et al. 2012; Colli et al. 2013). Recently **O**3 149 Bertolini et al. (2015) used a whole genome sequencing 150 approach to study the evolutionary aspects of Italian donkey 151 populations and to highlight divergence with horse genome. 152 In the present study, we analyzed the genetic variability at

in the present study, we analyzed the genetic variability at mitochondrial DNA level of six Italian autochthonous breeds. In particular, we considered two breeds living in Sardinia, Asinara and Sardo donkey, two breeds living in North Central Italy, Amiata donkey bred on the slopes of the Amiata mountain (Central Italy) and Romagnolo donkey bred in Emilia Romagna region (North Italy), and two breeds bred in South Italy, Martina Franca donkey and Ragusano donkey autochthonous of Puglia and Sicily respectively.

The aims of this study were: (i) to analyze genetic variability between six Italian donkey breeds, by means of mitochondrial DNA (mtDNA) D-loop sequences, (ii) to evaluate the presence of genetic breed sub-structuring and (iii) to compare mtDNA of Italian breeds with Modern European donkey breeds, African wild asses and Ethiopian donkey to understand their genetic relationships.

Materials and methods

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

Sample collection, PCR conditions, amplification, and sequencing protocols

174A total of 104 donkeys blood samples belonging to six Italian175breeds (Amiata donkey AMD=8; Asinara donkey ASD=28;176Ragusano donkey RAD=6; Romagnolo donkey ROD=8;177Martina Franca donkey MFD=5 and Sardo donkey SAD=49)

were available in an existing historical bio-bank at the Department of Veterinary Medicine of the University of Milan.

The mtDNA was isolated from whole blood samples (collected in 0.5 M EDTA and stored at -20 °C) using the NucleoSpin[®] Blood kit (Macherey-Nagel) following the manufacturer's instructions.

According to the complete donkey mtDNA sequence GenBank X97337 (Xu et al. 1996) two pairs of primers were designed to amplify a 478 bp fragment of the donkey mtDNA comprised between sites 15,386 to 15,863, using Primer3web version 4.0.0 (http://bioinfo.ut.ee/primer3) (F_DONK: CCCCAAGGACTATCAA for forward primer and R_DONK: GTTCTTTCTTCAGGGCCATT for reverse primer).

PCR were carried out using a MJ Research PTC-200 following conditions previously reported in Cozzi et al. (2004). PCR products were purified using SEPHADEX[®] G-50 (SIGMA) and sequenced for both strands using Big Dye[®] Terminator v 3.1 Cycle Sequencing Kit (Applied BiosystemsTM) according to the manufacturer's instructions. The PCR products were analyzed by ABI PRISM[®] 310 Genetic Analyser and the nucleotide composition was determined using specific software. The raw sequence trace files were checked for the presence of ambiguous bases using software Chromas v. 2.5.1 (http:// www.technelysium.com.au/). The donkey mtDNA sequences obtained were deposited in the GenBank (GenBank ID: KX622700-KX622727).

Data analyses

All the obtained sequences were aligned with the complete donkey mtDNA sequence using ClustalW algorithm implemented in BioEdit (Hall 1999) and MEGA v.7.0.14 software (Kumar et al. 2016).

Italian donkey population analyses

The maximum composite likelihood estimate of the nucleotide substitution pattern was calculated using MEGA v.7.0.14 software.

Number of polymorphic sites (S), parsimony informative (S_{Pl}) and singleton site (S_S) , number of haplotypes (NH), private (H_P) and shared haplotypes (H_S) , haplotype diversity (hd), nucleotide diversity (π) and average number of nucleotide differences (k) were calculated according to Tajima (1983) and Nei (1987) using DnaSP5 v.5.10.01 software (Librado & Rozas 2009). Pairwise population genetic differentiation (F_{ST}) and the average number of pairwise differences within and between populations were calculated using Arlequin v.3.5.2.2 (Excoffier & Lischer 2010).

The MEGA v.7.0.14 software was used to analyze the relationships among the haplotypes identified in Italian donkey populations and those identified in wild asses from GenBank (*Equus asinus somalicus, Equus asinus africanus* and *Equus hemionus luteus*) (GenBank ID are listed in the Supplementary Table S1).

Genetic relationships among breeds were reconstructed using Median-Joining Networks (MJN) by Network v.4.6 software (Bandelt et al. 1999). A Neighbour-joining (NJ) tree (Saitou & Nei 1987) was constructed based on the 232

233

234

235

236

178

179

180

Kimura-2-parameter model distances (Kimura 1980) calculated among the haplotypes using MEGA v.7.0.14 software. The bootstrap analysis, running 1000 bootstrap replicates, was applied in order to check the robustness of the resulting dendrogram. In addition, the Arlequin v.3.5.2.2 software was used to perform an Analysis of MOlecular VAriance (AMOVA).

Italian donkey breeds vs. other donkey populations The Italian mtDNA sequences were compared with 172 publically available domestic donkey's mtDNA D-loop sequences deposited in GenBank (http://www.ncbi.nlm.nih.gov/), comprising European domestic donkeys, wild asses and Ethiopian donkeys (GenBank ID are listed in the Supplementary Table S1).

All the 172 D-loop sequences were trimmed (236 bp overlapping positions) in order to allow the alignment and the comparison with Italian donkey breed sequences.

The DnaSP5 v.5.10.01 software was used in order to obtain haplotypes from all final D-loop sequences dataset. Based on the haplotypes a MJN by Network v.4.6 software (Bandelt et al. 1999) was constructed. Finally, the AMOVA analysis was performed using the final D-loop sequences dataset.

Results and discussion

Italian donkey population analyses

The mtDNA D-loop sequences were obtained for all analyzed samples. Twenty parsimony informative polymorphic sites were identified in the 104 Italian donkey mtDNA D-loop sequences 478 bp. According to the literature (Kumar et al. 2016) the parsimony informative sites are defined as mutations that have a minimum of two nucleotides that are present at least twice in the sampled population, whereas noninformative sites are singleton sites. All the polymorphic site were transition. The maximum composite likelihood estimate of the nucleotide substitution pattern, identified for the 104 sequences, was 30.55% (A), 27.1% (T), 29.1% (C), and 13.2% (G). Positions containing gaps and missing data were eliminated.

At population level, all the breeds (except for ROD) showed one or more singleton sites. As reported in Table 1 the MFD breed counted the highest number of non-informative singleton sites (S_s), as well as the lowest number of polymorphic sites (S) was identified in RAD breed.

In addition, the haplotypes identified in the analyzed breeds ranged from 3 (MFD) to 6 (AMD and SAD).

The breed genetic diversity (Table 1) evaluated by the nucleotide diversity value (π) ranged from 0.005 (RAD) to

0.025 (MFD) with an average value of 0.018. The haplotype diversity value (*hd*) ranged from 0.638 (ASD) and 0.893 (AMD) with an average value of 0.862. In addition, the average number of nucleotide differences (*k*) ranged from 1.2 (MFD) to 9.750 (ROD) with an average value of 8.412.

The MFD breed showed three haplotypes, the highest π value (0.02) and the highest number of parsimony sites (S_S = 11), but the haplotypes with the lowest nucleotide diversity (k = 1.2). On the contrary RAD have the lowest π (0.005) and K (k = 2.2) value, but haplotypes have a high hd value (0.867). The high k value (9.750) for the ROD breed is due to the high difference among the five haplotypes identified in the population.

Despite the low number of samples, a high *hd* value, with two private haplotypes (H6 and H11) for the AMD breed was found (Table 2). The two Sardinian breeds showed the presence of mitochondrial genetic variability, with similar *hd*, *k* and π values among them. These results confirm the findings obtained by Cosseddu et al. (2001) who used microsatellites markers to evaluate genetic variability between the two Sardinian breeds.

As reported in Table 1, all the average molecular indices show the presence of genetic variability among the Italian breeds. These results are similar to those observed in Croatian and Balkan donkeys by Ivankovic et al. (2002) and Pérez-Pardal et al. (2014), in Chinese donkeys by Lei et al. (2007) and Han et al. (2014), Ethiopian donkey by Kefena et al. (2014) and in Anatolian and Cypriot donkey populations by Cinar Kul et al. (2016).

Among these, the haplotype H14, identified in four donkeys, showed 18 nucleotide substitutions compared to the sequence X97337, whereas the haplotype H1 was identical to the reference sequence (Table 2).

The fifteen haplotypes found comparing our 104 donkey sequences with the donkey reference (accession number X97337) are shown in Table 2. All the haplotypes were shared by at least two breeds, except for the H6 (AMD), H8 (SAD), H10 (MFD), H11 (AMD) and H15 (SAD) haplotypes that are private ones (Table 2). Table 2 also reports the absolute and the relative haplotype frequencies. The most frequent haplotype found among the breeds was H3 (25%) followed by H2 (19.2%) and then H13 (17.3%), whereas H6 and H11 (1%) were the less frequent ones.

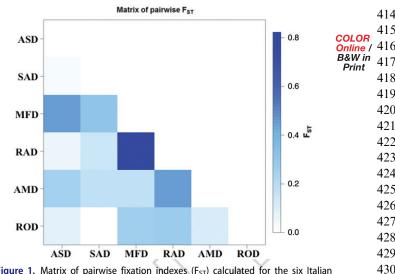
The matrix of pairwise fixation indexes (F_{ST}) is shown in Figure 1. The F_{ST} values range from 0.001 (SAD vs. ROD) to 0.822 (RAD vs. MFD). Except for AMD, the larger differences in F_{ST} values were found between the MFD breed and the

Table 1. D-loop nucleotide polymorphisms and molecular diversity indices of the six Italian donkey breeds.

Breed	n	S	S _{PI}	Ss	NH	P _H	S _H	$\pi \pm s.d.$	$hd \pm s.d.$	k
AMD	8	13	12	1	6	2	4	0.013 ± 0.0026	0.893 ± 0.012	6.143
ASD	28	18	17	1	4	0	4	0.015 ± 0.0025	0.638 ± 0.069	7.024
MFD	5	13	2	11	3	1	2	0.025 ± 0.0010	0.700 ± 0.048	1.2
RAD	6	5	4	1	4	0	4	0.005 ± 0.0009	0.867 ± 0.017	2.2
ROD	8	18	18	0	5	0	5	0.020 ± 0.0032	0.786 ± 0.151	9.75
SAD	49	20	19	1	6	2	4	0.019 ± 0.0010	0.711 ± 0.048	8.961

Sample size (*n*), total polymorphic sites (S), parsimony informative (S_{Pl}) and singleton site (S_{S}), number of haplotypes (NH), private haplotype (P_{H}), shared haplotype (S_{H}), nucleotide diversity (π), haplotype diversity (*hd*) with their standard deviations (s.d.) and average number of nucleotide differences (*k*) within and across the six populations.

	% Of haplotype frequency		1.9	19.2	25.0	4.8	1 9	01	2.9	2.9	3.8	2.9	1.0	1.9	17.3	4.8	8.7
Number of samples/breed	, 15,621 15,644 15,652 15,662 15,698 15,770 15,801 15,806 15,821 15,822 15,829 AMD ASD MFD RAD ROD SAD		-	1 1 3	23	2	-	-	1 2	m					10	4 1	6
Number of s	D ASD MFD			15	£	2					-	£		-	80		
	5,829 AMI	ט	-			-					۰ د			-	A	A	A
	15,822 15	ט			A									A	A	A	A
	15,821	U						•				г	F	г	⊢	н	⊢
	I 15,806	υ		•	•			•			•	•		•	⊢	⊢	⊢
	70 15,80	U		•	•			· ⊢	·	F	⊢	F	⊢	⊢	⊢	L	⊢
	598 15,77	-		•	•	•		•					_		U _	U _	U _
	5,662 15,	• •						•			ט	ט	ق	ט	ق	ט	ט
of nucleotide substitutions	15,652 1	U										г		г	Н	-	E
otide sub:	15,644	ט						•	. A	A	A	A	A	A	A	A	A
		A					2	-		-	9	6	9	U	9	9)
Position	98 15,59	A	•	U	5	9	0		2	9	9	9	9	J	U	5	U
	,597 15,5	E E	1	-)		-	-	5	>	⊢	۲	н		⊢	⊢	⊢
	5,590 15,	A	1	1	<												ט
	15,580 1	A						•			ט	ט	ט	ט		ט	ט
	15,569	A						ى .	, .	ט	ט	J	ט	ט	ט	ט	ט
	30 15,50 3	⊢		•	•			•			U	U	U	U	U	U	U
	484 15,45	0 0				•		•		•	•	•		•	A T	A T	A
	Haplotypes 15,484 15,490 15,503 15,569 15,580 15,590 15,597 15,598 15,599		•			•					•			•	4	4	4
	Hapl	X97337	H	H2	H3	H4	нv	He	H7	H8	6H	H10	H11	H12	H13	H14	H15





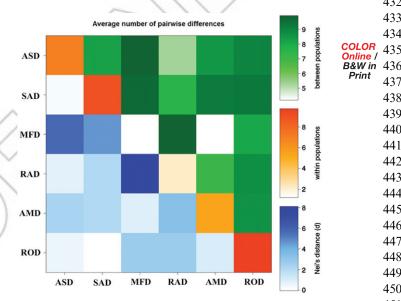


Figure 2. Nei's average number of pairwise differences (π) within and between the six Italian donkey breeds. On diagonal, the π within sampled populations in orange; above the diagonal, the π_{XY} between pairs of populations in green; below the diagonal, the Nei DA distance or the net number of nucleotide differences between the populations in blue.

other populations, ranging from 0.257 (MFD vs. ROD) to 0.822 (MFD vs. RAD).

The Nei's average number of pairwise differences within and between the populations calculated for the six Italian donkey breeds is shown in Figure 2.

MFD appears more separated than the other breeds showing the minor number of differences within samples in the population (the white box in orange diagonal in Figure 2). This may be due to conservation programs undergoing for a long time in this breed aiming in maintaining the breed peculiarities without crossing with non MFD animals (Rizzi et al. 2011).

The two Sardinian breeds (ASD and SAD) showed a high within population nucleotide variability (Figure 2). The low F_{ST} value (Figure 1) found between the two Sardinian populations (ASD vs. SAD = 0.031) highlights their closely genetic

Q6

relationship at mtDNA level, despite the evident differences 473 in morphological characteristics (i.e. coat colour) of these 474 breeds. In fact, one of the hypotheses on the origin of ASD is 475 based on the possibility that the albino mutation occurred in 476 a 'grey' haplotype, present in the Sardinia donkey. It is highly 477 probable that the isolation of the ASD donkey population 478 and its high-inbreeding level have raised the frequency of 479 the 'albino pattern' in this population (Utzeri et al. 2016). The 480 closely relationship between the two populations was found 481 by Colli et al. (2013) who reported a similar F_{ST} value (0.041) 482 using a panel of 16 microsatellites. The same result holds 483 using the Nei's distance as shown in Figure 2. 484

The F_{ST} values (Figure 1) show the closely relationship 485 between SAD and ROD (SAD vs. ROD = 0.001) and SAD and 486 RAD (SAD vs. RAD = 0.152), but they well differentiated SAD 487 from MFD (SAD vs. MFD = 0.319) as well as the Nei's distan-488 ces (Figure 2) do. In fact, in the last century, SAD probably 489 490 received some genetic contributions from other populations 491 including MFD, RAD and ROD, in order to increase its structural dimension even if, in the last decade the Dipartimento 492 di Ricerca per l'Incremento Ippico (AGRIS-Sardegna) has 493 implemented the programs for the SAD breed conservation 494 (Cherchi 2005). 495

The ROD breed is close to RAD (ROD vs. RAD = 0.279), 496 MFD (ROD vs. MFD = 0.257) and AMD (ROD vs. AMD = 0.109), 497 confirming the historical information indicating that these 498 three breeds contributed to the genetic makeup of the ROD 499 gene pool (Beretti et al. 2005; Salza 2006; Kugler et al. 500 2007–2008) (Figure 1). In fact, ROD showed the highest gen-501 etic variability within population (in diagonal the darkest 502 orange box in Figure 2). 503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

The AMD breed, characterized by morphological ancestral features referable to wild asses (stripped limbs, the dark dorsal strip and mouse coat colour), appeared closely related to all the populations (Figure 1) except for RAD (AMD vs. RAD = 0.458) and showed a medium nucleotide variability value within breed (in diagonal the orange box in Figure 2). Unlike what reported by Colli et al. (2013), our data high-lighted a high haplotypes variability of the AMD mtDNA, showing four shared haplotypes and two private haplotypes.

Also the NJ dendrogram in Figure 3, constructed using our haplotypes and mtDNA sequences from *Equus asinus africanus* (Nubian wild ass), *Equus asinus somalicus* (Somali wild ass), and *Equus hemionus luteus* sequence (used as outgroup), did not confirm the similarity between the mtDNA sequences of AMD donkeys and *Equus asinus somalicus*, as reported by Colli et al. (2013). In fact, in the dendrogram the *Equus asinus somalicus* cluster is clearly separated from the one including both ours and *Equus asinus africanus* sequences. The bootstrap values at the third nodes (95), indicating the robustness of the dendrogram, support the clustering structure outcome in Figure 3.

Regarding the RAD we found four haplotypes with high similarity among them resulting in a low within population variability (in diagonal the light orange box in Figure 2). These findings do not agree with those found by Bordonaro et al. (2012) using 14 microsatellites markers.

The Figure 4(a) shows the MJN constructed based on the 104 Italian donkey mtDNA sequences and the reference

sequence X97337. Three haplogroups (A, B and C), separated from 5 nucleotide substitutions each, were found. In detail, the three haplogroups include haplotypes that differ each other by few mutations: 1 to 5 mutations for haplogroup A; 1 to 3 mutations for haplogroup B; 2 to 3 mutations for haplogroup C. In addition, another haplotypes NJ tree, based on Kimura-2-parameter model distances, allowed us to confirm the haplogroups clustering distribution (Figure 4(b)). These haplogroups comprise haplotypes for which common origins are assumed since they share a characteristic pattern of mutations.

The AMOVA analysis is a useful tool to check how the genetic diversity is distributed among populations, whose structure is quantified by F_{ST} . We examined different possible structures by creating and comparing different population groups. We ran the analysis under two hypotheses:

Hypothesis (1) three groups according to the geographical prevalence, i.e. group 1, ASD and SAD (Sardinian populations), group 2, AMD and ROD (North Central Italy) and group 3, MFD and RAD (South Italy);

Hypothesis 2) three groups according to the donkey height at the withers, i.e. group 1, ASD and SAD (height at the withers low), group 2, MFD, ROD and RAD (height at the withers high) and group 3, AMD (height at the withers intermediate).

The Table 3 reports the results for the AMOVA analysis according to two hypotheses. The results highlight that the majority of the observed variance is due to differences among samples within breeds. The most part of the variation is observed within the populations (85.5% Hypothesis 1 and 85.2% Hypothesis 2), whereas the differences among groups represent only the 1.5% and 2.18% of the variation respectively (Table 3).

Italian donkey breeds vs. other donkey populations

The Italian mtDNA sequences were compared with 172 publically available mtDNA D-loop sequences belonging to European domestic and Ethiopian donkeys and wild asses. In addition, to assess a hierarchical structure among Italian populations, European domestic donkeys, Ethiopian donkeys and wild asses, the AMOVA analysis was performed under different hypotheses grouping breeds in different clusters. The hypothesis that better defined the clustering strategy in groups resulted to be in agreement with the historical data on the of Italian donkey breeds origin. According to this hypothesis the breeds have been clustered in six groups: (1) Equus asinus africanus (Nubian wild ass), Equus asinus somalicus (Somali wild ass); (2) Ethiopian donkeys and AMD; (3) ASD and SAD; (4) Spanish donkey, MFD, RAD and ROD; (5) Balkan donkey; (6) Turkey and Cyprus donkey breeds (Table 4). As expected the variation within population of 54.02%, highlights the differences among samples within breeds (Table 4). A considerable amount of variation (37.64%) can be attributed to differences found among populations within defined groups. Instead, the variance among groups resulted low (8.34%) (Table 4)

The MJN analysis identified 62 haplotypes from 276 mtDNA D-loop sequences (172 GenBank sequences and our

590

532

533

534

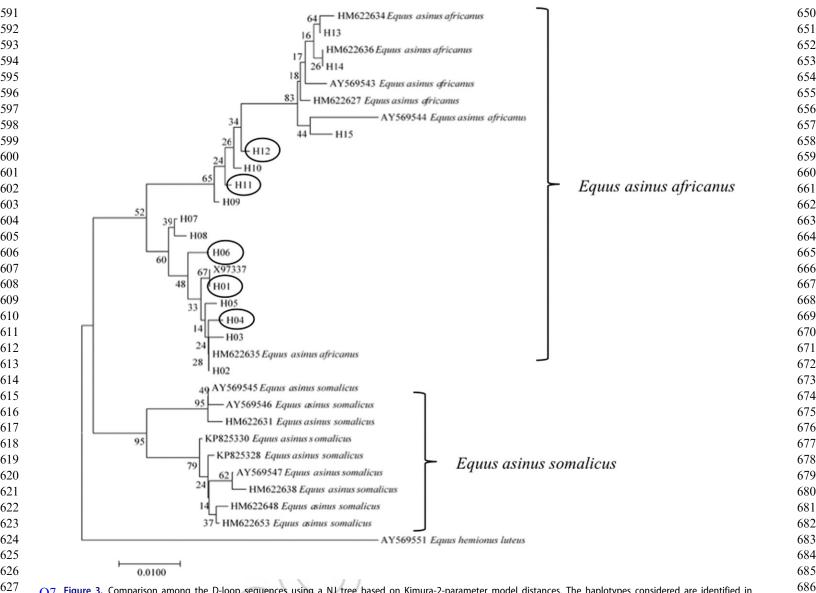


Figure 3. Comparison among the D-loop sequences using a NJ tree based on Kimura-2-parameter model distances. The haplotypes considered are identified in *Equus asinus africanus* (Nubian wild ass), *Equus asinus somalicus* (Somali wild ass), *Equus hemionus luteus* (outgroup) and Italian donkey breeds. The bootstrap values represent the robustness of the dendrogram. In black circle, the haplotypes with Amiata donkey D-loop sequences.

630 104 Italian donkey breeds) showing a high variability 631 (Hd = 0.931) (Supplementary Table S2). In Figure 5 two differ-632 ent clades are clearly defined: clade 1 includes the 61,6% of 633 sequences, whereas clade 2 includes the 38,4% of sequences. 634 Out of 62 haplotypes, 58 are referred to the Nubian lineage, 635 whereas only four haplotypes belong to Somali lineage. The 636 Somali haplotypes are separated by clade 1 and 2 by 6 muta-637 tions (15,489, 15,536, 15,541, 15,637, 15,652 and 15,704) that 638 represent a specific 'pattern' of Equus asinus somalicus, except 639 for the 15,652 mutation also present in Equus asinus africanus 640 (Figure 5). Our results are consistent with those obtained by 641 Kimura et al. (2011) who identified two distinct clades sepa-642 rated from the Equus asinus somalicus group by 12 643 644 mutations.

645The analysis clearly defines four main represented haplo-646types: H2 (74%), H3 (42%), H5 (32%) and H6 (61%). Despite647previous published evidences, our sequences appeared to be648related only with Equus asinus africanus (Nubian wild ass),649and then we did not confirm the relationship between

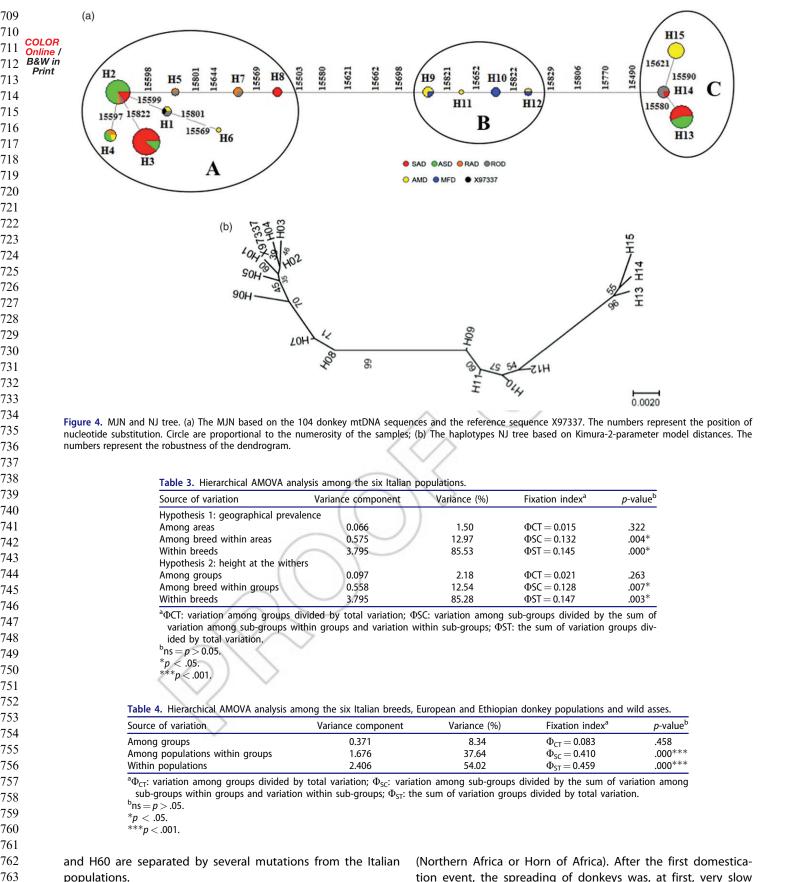
Amiata donkey and *Equus asinus somalicus* (Somali wild ass) as cited in works by other authors (Colli et al. 2013) as a personal communication.

The influence of Spanish donkey populations, particularly the Catalana breed on MFD, is underlined by the common haplotype H1 and the closely relationships between H10 vs. H29 and H10 vs. H30 separated by mutation 15,592 and 15,621, respectively.

The ASD samples and the most of SAD ones are grouped in three haplotypes (H2, H3 and H5). RAD and ROD shared haplotypes H11 and H12, thus confirming the closely relationships here in before discussed. These two haplotypes derived from H2 that includes the most common haplotype found in Spanish donkeys (Aranguren-Mendez et al. 2004), confirming once again the closely relationships between the two Italian breeds and the Spanish donkeys.

The haplotypes H2 and H6 comprise the most part of Balkan donkeys and the Italian donkey populations. The Anatolian and Cyprus donkey haplotypes H55, H56, H57, H59 705

706



Pérez-Pardal et al. (2014) reported similar results and
underlined the hypothesis that the donkey mtDNA genetic
background probably results from contemporaneous domestication events occurred in a limited geographical area

(Northern Africa or Horn of Africa). After the first domestication event, the spreading of donkeys was, at first, very slow and scattered. The importance of the donkey's diffusion can be attributed to the evolution of long-distance trade and to the human migrations, especially occurred in Africa (Blench 2000). Afterwards, the Roman played an important role in

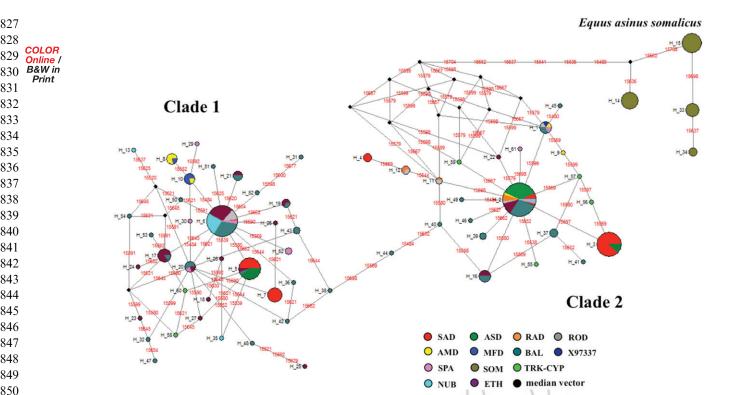


Figure 5. MJN based on the six Italian donkey breeds and the 172 donkey sequences from GenBank. Circle are proportional to the haplotypes frequencies; black circles are median vectors (mv) representing extant unsampled or extinct ancestral sequences; in red the number of mutation point respect to the reference sequence X97337. SAD = Sardo donkey; ASD = Asinara donkey; AMD = Amiata donkey; RAD = Ragusano donkey; ROD = Romagnolo donkey; MFD = Martina Franca donkey; BAL = Balcan donkey; SPA = Spanish donkey; SOM = Equus asinus somalicus; TRK-CYP = Turkey and Cyprus donkey; NUB = Equus asinus somalicus; ETH = Ethiopian donkey. GenBank ID are listed in the S1 supplementary table.

expansion of donkeys in neighbouring territories. In fact, the presence of donkeys belonging to clade 1 and 2 is recorded by archaeological samples from Pompei (Kimura et al. 2013).

Finally, it is interesting to note that at the beginning of the twenty-first century the trend of expansion of donkey populations, occurred in rural areas of African countries, Asia and Latin America, where human populations did not have access to mechanical power for agricultural practices. On the contrary in the industrialized countries, despite the use of donkeys for recreational purposes such as ecotourism, trekking, equestrian therapy/donkey therapy (especially with children) and for the production of milk to be used in cosmetics industry, or for human consumption in cases of cow milk allergy, the donkey populations remain relatively stable (Starkey& Starkey 2000).

Conclusions

Despite the relatively high number of references on genetic variability of Italian donkey breeds analyzed by microsatellite markers and morphological traits, this is the first study on mtDNA D-loop genetic variability of Italian donkey breeds.

The results of our study show the existing genetic variabil-ity within and between breeds found in the six donkey popu-lations bred in Italy. However, a geographical clustering among Italian populations was not found. The identified hap-lotypes were grouped in three haplogroups, each differenti-ated from the others by few mutations.

The relationships among our Italian donkey breed sequences and those from donkeys belonging to breeds living in Mediterranean and Balkan areas, showed the complexity of the ancestry and of the genetic makeup of the modern donkey populations.

To preserve the genetic resources represented by local populations will be one of the main goal for the future generations. The maintenance of the donkey biodiversity, nowadays at risk of loss if not of extinction, requires comprehensive knowledge of the donkey breeds characteristics, including data on population size and structure, breed geographical distribution and the within and between breed genetic diversity assessed by different marker types, including mtDNA sequences. Assessment of genetic variation as done in this study represents one of the indication that can be used in ex situ genetic conservation, helping the prioritization of individuals to be collected and stored in germplasm cryobanks.

Acknowledgments

The authors thank Dipartimento di Ricerca per l'Incremento Ippico (AGRIS-Sardegna) and staff who provided the samples.

Disclosure statement

The authors declare they have no competing interests.

Funding

This study was funded by internal funds.

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

1057

1058

1059

1060

945 References

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

984

985

- 946 Aranguren-Mendez J, Beja-Pereira A, Avellanet R, Dzama K, Jordana J. 947 2004. Mitochondrial DNA variation and genetic relationships in 948 Spanish donkey breeds (Equus asinus). J An Breed Genet. 949 121:319-330.
- Bandelt H-J, Forster P, Rohl A. 1999. Median-Joining networks for infer-950 ring intraspecific phylogenies. Mol Biol Evol. 16: 37-48. 951
 - Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhiet AO, Abdalla MA, Mashkour M, Jordana J, Taberlet P, Luikart G. 2004. African origins of the domestic donkey. Science. 304:1781.
 - Beretti V, Zanon A, Soffiantini CS, Sabbioni A. 2005. Preliminary results about morphological and demographic traits of Romagnolo donkey. Annali Della Facoltà Di Medicina Veterinaria Di Parma. 25:131-144.
 - Bertolini F, Scimone C, Geraci C, Schiavo G, Utzeri VJ, Chiofalo V, Fontanesi L. 2015. Next Generation semiconductor based sequencing of the donkey (Equus asinus) genome provided comparative sequence data against the horse genome and a few millions of single nucleotide polymorphisms. PLoS One. 10:e0131925.
 - Blench RM. 2000. A history of donkeys, wild asses and mules in Africa. In: Blench RM, & MacDonald KC, editors. The origins and development of African livestock: Archaeology, genetics, linguistics and ethnography. London: UCL; p. 339-354.
 - Bordonaro S, Guastella AM, Criscione A, Zuccaro A, Marletta D. 2012. Genetic diversity and variability in endangered Pantesco and two other Sicilian donkey breeds assessed by microsatellite markers. Sci World J. (Article ID 648427, 6 pages.) 04
 - Chen SY, Zhou F, Xiao H, Sha T, Wu SF, Zhang P. 2006. Mitochondrial DNA diversity and population structure of four Chinese donkey breeds. Anim Genet. 37:427-429.
 - Cherchi R. 2005. Asino Sardo [Sardinian donkey]. Paper presented at: 1° Convegno Nazionale sull'asino. 28-29 May 2005, Grosseto, Italy. (Italian)
 - Ciampolini R, Cecchi F, Mazzanti E, Ciani E, Tancredi M, De Sanctis B. 2007. The genetic variability analysis of the Amiata donkey breed by molecular data. Italian J Anim Sci. 6(Suppl 1):78-80.
 - Cinar Kul B, Bilgen N, Akyuz B, Ertugrul O. 2016. Molecular phylogeny of Anatolian and Cypriot donkey populations based on mitochondrial DNA and Y-chromosomal STRs. Ankara Üniv Vet Fak Derg. 63:143–149.
 - Colli L, Perrotta G, Negrini R, Bomba L, Bigi D, Zambonelli P, Verini Supplizi A, Liotta L, Ajmone-Marsan P. 2013. Detecting population structure and recent demographic history in endangered livestock breeds: the case of the Italian autochthonous donkeys. Anim Genet. 44:69-78
- Cosseddu GM, Fraghì A, Mura L, Carta A, Cherchi R, Pau S. 2001. Relazioni genetiche tra le popolazioni asinine della Sardegna. Analisi 983 con marcatori molecolari [Genetic relationships among donkey populations living in Sardinia. An analysis using molecular markers]. Ippologia. 12:25-33. (Italian).
- 986 Cozzi MC, Strillacci MG, Valiati P, Bighignoli B, Cancedda M, Zanotti M. 987 2004. Mitochondrial D-loop sequence variation among Italian horse breeds. Genet Sel Evol. 36:663-672. 988
- Cristofalo C, Cancadda M, Cozzi MC, Valiati P. 1994. Salvaguardia della 989 variabilità genetica: analisi del polimorfismo ematico di una popola-990 zione di asini sardi. [Conservation of genetic variability: analysis of 991 blood polymorphism of a population of Sardinian donkeys] Atti XLVIII 992 Conv. SISVET. 48:1821-1825. (Italian).
- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of pro-993 grams to perform population genetics analyses under Linux and 994 Windows. Mol Ecol Resources. 10:564-567. 995
- Guastella AM, Zuccaro A, Bordonaro S, Criscione A, Marletta D, D'Urso G. 996 2007. Genetic diversity and relationship among the three autochthon-997 ous Sicilian donkey populations assessed by microsatellite markers. 998 Italian J Anim Sci. 6(Suppl 1):143.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment edi-999 tor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 1000 41:95-98 1001
- Han L, Zhu S, Ning C, Cai D, Wang K, Chen Q, Hu S, Yang J, Shao J, Zhu 1002 H, et al. 2014. Ancient DNA provides new insight into the maternal lin-1003 eages and domestication of Chinese donkeys. BMC Evol Biol. 14:246.

- Ivankovic A, Kavar T, Caput P, Mioc B, Pavic V, Dovc P. 2002. Genetic diversity of three donkey populations in the Croatian coastal region. Anim Genet. 33:169-177
- Kefena E, Dessie T, Tegegne A, Beja-Pereira A, Yusuf Kurtu M, Rosenbom S, Han JL. 2014. Genetic diversity and matrilineal genetic signature of native Ethiopian donkeys (Equus asinus) inferred from mitochondrial DNA sequence polymorphism. Livestock Sci. 167:73-79.
- Kimura B, Marshall F, Beja-Pereira A, Mulligan C. 2013. Donkey Domestication, Afr Archaeol Rev. 30:83-95.
- Kimura B, Marshall FB, Chen S, Rosenbom S, Moehlman PD, Tuross N, Sabin RC, Peters J, Barich B, Yohannes H, et al. 2011. Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. Proc Biol Sci. 278:50-57.
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16:111-120.
- Kugler W, Grunenfelder HP, Broxham E. 2007-2008. Donkey breeds in Europe. Inventory, Description, Need for Action, Conservation. [Internet]. Report 2007-2008. [Internet]. Available from: http://www. boerenvee.nl/uploads/ezelstudie.pdf
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. 33:1870-1874.
- Lei C, Ge Q, Zhang H, Liu R, Zhang W, Jiang Y, Dang R, Zheng H, Hou W, Chen H. 2007. African maternal origin and genetic diversity of Chinese domestic donkeys. Asian Aust J Anim Sci. 20:645-652.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25:1451-1452.
- Lopez LC, Alonso R, de Aluja AS. 2005. Study of the genetic origin of the Mexican creole donkey (Equus asinus) by means of the analysis of the D-Loop region of mitochondrial DNA. Trop Anim Health Prod. 37(Suppl1):173-188
- Marshall F. 2007. African pastoral perspectives on domestication of the donkey. In: Denham T, Iriarte J, & Vrydaghs L, editors. Rethinking agriculture: Archaeological and ethnoarchaeological perspectives. Walnut Creek: Left Coast Press; p. 371-407.
- Matassino D, Cecchi F, Ciani F, Incoronato C, Occidente M, Santoro L, Ciampolini R. 2014. Genetic diversity and variability in two Italian autochthonous donkey genetic types assessed by microsatellite markers. Italian. J Anim Sci. 13:3028.
- Moehlman PD. 2002. Equids: Zebras, Asses and Horses. Status Survey and Conservation Action Plan. IUCN/SSC Equid Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK.
- Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.
- Oakenfull EA, Lim HN, Ryder OA. 2000. A survey of equid mitochondrial DNA: Implications for the evolution, genetic diversity and conservation of Equus. Conserv Genet. 1:341-355.
- Pérez-Pardal L, Grizelj J, Traoré A, Cubric-Curik V, Arsenos G, Dovenski T, Marković B, Fernández I, Cuervo M, Alvarez I, et al. 2014. Lack of mitochondrial DNA structure in Balkan donkey is consistent with a quick spread of the species after domestication. Anim Genet. 45:144-147.
- Rizzi R, Tullo E, Cito AM, Caroli A, Pieragostini E. 2011. Monitoring of genetic diversity in the endangered Martina Franca donkey population. J Anim Sci. 89:1304-1311.
- Rossel S, Marshall F, Peters J, Pilgram T, Adams MD, O'Connor D. 2008. Domestication of the donkey: Timing, processes, and indicators. Proc Natl Acad Sci USA. 105:3715-3720.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406-425.
- Salza G. 2006. L'asino Romagnolo, in sella tra biodiversità e tempo libero. [The Romagnolo donkey, riding between biodiversity and recreation]. Agricoltura. 34:18–20. (Italian).
- Starkey P, Starkey M. 2000. Regional and world trends in donkey populations. In: Starkey P, Fielding D, editors. Donkeys, People and Development. A resource book of the Animal Traction Network for Eastern and Southern Africa (ATNESA), ACP-EU Technical Centre for Agricultural and Rural Cooperation (CTA), Wageningen, The Netherlands; p. 247. Available from: http://www.atnesa.org.

1061 1062 **O**5

 Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics. 105:437–460.

Utzeri VJ, Bertolini F, Ribani A, Schiavo G, Dall'Olio S, Fontanesi L. 2016. The albinism of the feral Asinara white donkeys (Equus asinus) is determined by a missense mutation in a highly conserved position of the tyrosinase (TYR) gene deduced protein. Anim Genet. 47:120–124.

Xu X, Gullberg A, Arnason U. 1996. The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. J Mol Evol. 43:438–446.