

## Developing a microsatellite markers-based parentage test for *Camelus dromedarius*

G. NOUAÏRIA<sup>1\*</sup>, M. CASSANDRO<sup>2</sup>, M. BEN LARBI<sup>3</sup>, W. FENNANI<sup>3</sup>, E. VIALE<sup>2</sup>, M. HAMMADI<sup>1</sup>, M.H. YAHYAOUTI<sup>1</sup>

<sup>1</sup> Livestock and Wildlife Laboratory. Arid Lands Institute. University of Gabes. Road El Jorf, 4119 Medenine, Tunisia

<sup>2</sup> Department of Agronomy, Food, Natural resources, Animals and Environment. University of Padova. 35020 Legnaro, Italy

<sup>3</sup> Research Unit of biodiversity and resource valorization in mountain areas (UR17AGR14). High School of Agriculture in Mateur. University of Carthage. Road of Tabarka, 7030 Mateur, Tunisia

\*Corresponding author: nouairiaghada@gmail.com

**Abstract** – Although Tunisian cameline population counts more than 80 thousands heads, individual identification of these animals is often absent. This is a limiting factor to the accurate selection and genetic improvement efforts. With an efficiency approaching 100 %, parentage testing by genetic markers is a reliable and concluding scientific means to overcome these issues.

The present work evaluates twenty microsatellites to develop a reliable paternity test in *Camelus dromedarius*. After testing the markers on a total of 130 animals, we estimated for each marker, polymorphism, pairwise linkage disequilibrium and Hardy Weinberg equilibrium. A panel of 38 samples presenting 17 trios (sire, dam, and their offspring) was used to evaluate different sets of microsatellites as parentage tests and optimize the number of markers for economic reasons while retaining an assignment rate exceeding 99 %. Finally, a list of 12 microsatellites proved higher than 99% of exclusion power when one parent is unknown parent and 98% when both parents are unknown, was proposed as an efficient and economic parentage test for *Camelus dromedarius*.

**Keywords:** Dromedary camel, paternity test, genealogy, microsatellites.

### 1. Introduction

The need for parentage testing has been perceived since the early development of molecular techniques, by the 1960s. At that time, the blood groups were used for parentage testing in more than a livestock species (Rendel, 1957; Stormont, 1967) with an exclusion power limited to 30% (Douris, 1922). Now, Short Tandem Repeats (STR) based parentage tests are available for almost all- livestock species as well as for a large number of wild animal species (Cattle: Heyen *et al.*, 1997; Goat: Luikart *et al.*, 1999; da Silva *et al.*, 2014; Sheep: Al-Atiyat, 2015). Microsatellites are deemed useful because their efficiency exceeds 99%, their cost is affordable by most livestock owners, and the techniques are customary in most laboratories. For dromedary camels, a Random Amplified Polymorphic DNA-based (RAPD) test was described for parentage testing in Saudi Arabia (Al Swailem *et al.*, 2008) and another based on microsatellite markers was developed for race dromedaries (Spencer *et al.*, 2010).

Camel breeding in Tunisia and other nearby countries is practiced mostly in a traditional way, individual animal identification is not systematically practiced and no herd book is currently available.

Furthermore, mating camels is not controlled, thus no selection could be envisaged on the animals. Camel breeding in Tunisia is evolving towards more intensive systems and characteristics such as high milk potential, hefty animals and docile character, for instance, are sought for. However, animal selection is currently hampered by the lack of reliable data, since camel herders in Tunisia usually do not keep records of the animal mating and performances. Camels are of high economic value in Tunisia; their price ranges from 1000 DNT, for young males destined to slaughtering, to 8000 DNT or even higher for those used in tourism activities. Thus, a conclusive low-cost technique to identify the animal is of utter interest.

In the present work, 20 microsatellites were tested on samples from the Tunisian camel population in order to evaluate their suitability and effectiveness for the development of a parentage test in this species.



## 2. Materials and Methods

### 2.1. Sampling

A total of 130 blood samples were collected from -- different herds in southern Tunisia. An additional 36 samples from the IRA's experimental herd were included, corresponding to 17 trios (2 sires, 17 distinct dams, and their 17 offspring). The trio samples were distributed over three calving seasons (2014 to 2016). The camel ought to be caught and completely immobilized in order to safely draw blood from the jugular vein in 5 ml vacutainer tubes containing EDTA (Ethylene-diamine-Tetra-acetic Acid) to prevent coagulation. Samples were conserved at -20°C for later use.

### 2.2. DNA Extraction

After sample thawing, 250 µl of blood were used to extract DNA following the salting-out protocol (Miller *et al.*, 1988). Quantity and quality control were performed by optical density (OD) measurements at two wave lengths; 260 nm and 280 nm, using an optical spectrophotometer. For samples with low purity ( $OD_{260}/OD_{280} < 1.8$ ), an additional purification by phenol-chloroform (Sambrook and Russel, 2001) was performed.

### 2.3. DNA Amplification

Twenty microsatellite markers (Mariasegaram *et al.*, 2000; Penedo *et al.*, 1999; Evdotchenko *et al.*, 2003; Lang *et al.*, 1996) were amplified in three multiplex PCR reactions; two containing 7 microsatellites each and one containing only 6. Table 1 details a description of the used microsatellite markers. In order to standardize the PCR protocol, all samples were diluted to a final concentration of 30 ng/µl. The PCR reactions were then conducted in 10 µl final volumes using the Qiagen multiplex PCR kit.

**Table 1.** Annealing temperature ( $T_a$ ), allele size range, and GenBank accession number for the microsatellite markers

Locus	$T_a$	Allele size range (bp)	Gene Bank Accession number
CMS9	60	225-256	AF329160
CMS121	60	147-173	AF329159
CMS13	58	236-265	AF329158
CMS15	58	116-144	AF329151
CMS17	58	135-167	AF329147
CMS18	60	157-170	AF329148
CMS25	58	90-110	AF380345
CMS32	58	198-209	AF329146
CMS50	60	129-190	AF329149
CVRL01	60	188-253	AF217601
CVRL04	60	156-180	AF329149
CVRL05	60	148-179	AF217602
CVRL06	60	155-175	AF217606
CVRL07	60	265-315	AF217607
CVRL08	60	191-209	AF329147
LCA66	60	212-262	AF091125
VOLP10	60	236-268	AF305237
VOLP32	58	256-290	AF305234
YWLL44	60	86-120	AF329160
YWLL59	60	103-117	-

The thermal profile included 35 cycles of denaturation (30 s) at 95 °C, annealing (90 s) at 60 °C for multiplexes I and II and 58 °C for multiplex III, and elongation at 72 °C (1 min), for each cycle. An initial denaturation of 2 min at 95 °C and a final elongation of 30 min at 60 °C were also applied. The PCR products were then stored at -20 °C.

### 2.4. Capillary electrophoresis

For the capillary electrophoresis, one microliter µl from each PCR product sample (a minimum of 10 ng/marker is required) was transferred in a 96-wells plate and dried at 65 °C for 4 min as recommended by the sequencing laboratory (BMR Genomics, <http://www.bmr-genomics.it>). If necessary, the fragment detection was repeated using 2 µl of the samples. The capillary electrophoresis was realized in an Applied Biosystems AB3500 Genetic Analyzer, using Genscan Liz 500 (Applied Biosystems). The output files were read and analyzed using PEAKSCANNER V1.0 software (Applied Biosystems).

## 2.5. Statistical analyses

The number of alleles, polymorphic information content (PIC), Hardy-Weinberg equilibrium, and linkage disequilibrium were estimated using PopGene version 1.31 (Yeh *et al.*, 1999). In addition, the pairwise relatedness and inbreeding were estimated for the trio samples and for the whole population by maximum likelihood statistics using Coancestry software version 1.0.2 (Wang, 2008).

## 2.6. Parentage test developing scheme

In order to obtain an efficient but economical final parentage test, different combinations of the microsatellite markers were tested. Markers with highest NA and PIC were tested first, as they are expected to be more informative. The software CERVUS 3.0 (Marshall *et al.*, 1998) was used to evaluate the exclusion power of each set of markers based on the likelihood ratio scores (LOD) and error rate and the assignment confidence for each tested pair of animals.

The experiment design consisted in two scenarios described hereafter and the evaluation of the efficiency of each marker set was based on the LOD, assignment rate and confidence estimates. In the first test, for each offspring, both the father and the mother were masked. CERVUS estimated the probability of each offspring-mother pair as well as offspring-father pair following the highest LOD. The second test consisted in inferring paternity with known maternal origin, which –in practice- is very common inquiry. For each offspring, we suppose the mother is declared and the father is one of the two sires. Considering the highest LOD, all the 17 offspring were assigned to a father. The accuracy of the marker set was then evaluated by calculating the error rate.

## 3. Results and discussion

### 3.1. Statistical Analyses

Initially, the 20 microsatellite markers were tested on 130 samples from different regions of southern Tunisia in order to answer basic criteria before evaluating their efficiency in parentage testing. We checked (i) the microsatellite polymorphism at a threshold  $p = 0.005$ , (ii) the Hardy-Weinberg equilibrium of each locus to eliminate any effect of evolutionary change that can affect its frequencies, and (iii) the pairwise linkage disequilibrium to eliminate any redundant information. Following the predetermined criteria, CMS17 was eliminated as it was monomorphic. All the other 19 microsatellites were polymorphic at 99 % threshold; the maximum major allele frequency was 0.8. In fact, CMS17 was developed for *C. bactrianus* and appears to be monomorphic in *C. dromedarius*, as Al Mathen *et al.* (2016) reported only one allele in this locus in samples from 20 countries. Locus CVRL6 was also eliminated as it was transmitted in linkage with CMS52 ( $p = 0.067$ ) and was highly correlated to 4 other loci with a correlation exceeding 75 % (Weir, 1979). All the remaining markers were in no significant linkage ( $p = 0.05$ ).

With the exception of the locus VOLP32 that was bi-allelic, all the remaining loci were multi-allelic (table 2). Regarding the allele count, CMS12, CVRL5, CVRL1, YWLL44 and CVRL7 had more than 10 alleles in the Tunisian camel population. VOLP32, CMS18 and CMS32 had less than 5 alleles (table 2). The Polymorphic Information Content (PIC) is relative to the number of alleles and their allele frequencies. PIC was higher than 70 % for most loci; CVRL1, CMS50, CMS15, CVRL7, VOLP10, LCA66, CVRL5, CMS121, CMS9 and CMS13 (table 2). Five loci had relatively low PIC; VOLP32 (0.2), CVRL6, and CMS18 (0.3). For economic reasons, we tried to restrain the number of markers in the parentage assay. Thus, markers with higher number of alleles and higher PIC were included first in the different tested combinations for parentage assays.

**Table 2.** Number of alleles and Polymorphic Information Content (PIC) for the studied loci

Locus	Number of Alleles	Polymorphic Information Content
CMS9	7	0.707
CMS13	8	0.773
CMS15	8	0.748
CMS18	3	0.347
CMS25	5	0.633
CMS32	4	0.63
CMS50	8	0.829
CMS121	11	0.749
CVRL1	17	0.808
CVRL4	5	0.619
CVRL5	10	0.701
CVRL6	3	0.371
CVRL7	10	0.79
CVRL8	5	0.401
LCA66	8	0.734
VOLP10	7	0.73
VOLP32	2	0.211
YWLL44	11	0.596
YWLL59	5	0.538
<b>Mean</b>	7.368	0.627
<b>SD</b>	5.31	0.176

### 3.2. Exclusion power for parentage test

In general, the Tunisian camel population exhibits a mean of 15% of inbreeding (Ould Ahmed *et al.*, 2010; Nouairia *et al.*, 2015). The developed test should then be efficient in an average inbred herd and even more efficient between individuals from distinct herds. To minimize the test cost; we tried to limit the number of markers in the parentage assay as low as possible. Table 3 details the exclusion power to detect one unknown parent or two unknown parents when all the markers are used, and when we reduce the number of markers. After testing different number of marker sets, we found that a minimum of 13 microsatellites were needed in the double unknown parents test to yield a likelihood (LOD) of 99 %. Only 10 markers were necessary if one parent is known, to reach a power exclusion of 99 %. The different sets of 12 markers yielding the best assignment rates (99% or higher) with the minimal error rate were then compared on an economical basis. Preferably, the markers should be run in one multiplex, i.e; have the same annealing temperature. Thus, the test could be run in one PCR reaction and one capillary sequencing operation.

**Table 3.** Exclusion power of some of the tested combinations of markers

Number of markers	Exclusion power	Exclusion power
	<i>two unknown parents</i>	<i>one unknown parent</i>
18 markers	99.99 %	99.99 %
12 markers	98.61 %	99.99 %
8 markers	72.82 %	93.43 %

A final parentage test consisting of 12 microsatellite markers was set. With a declared mother, the markers test was able to assign all the 17 offspring to one of two genitors with a likelihood of 99.92 % and an estimated error rate of  $10^{-9}$ . Compared to the real data, all the young animals were rightly assigned to their fathers with no error. In the double unknown parents assay, all the fathers were correctly assigned. The mothers, coming from the same herd and being related (34% of shared genetic material), were more difficult to assign. The correct mother had the highest LOD score for all the tested offspring (17). This score was of 98 % or higher. But, in less than 20 % (3 individuals), the second highest LOD score was higher than 97 %. There was no error in the 17 assigned offspring but the mean estimated error rate based on CERVUS calculations was  $10^{-3}$ .

The cost of the parentage test was estimated to 40 DNT per individual; it includes all the reagents and material fees in addition to manipulation fees.

#### 4. Conclusion

The developed (proposed) parentage test uses 12 of the most informative microsatellite markers with an exclusion power of 99 % when one parent is known and 98 % if both parents are unknown. The test is based on techniques that are customary in most research laboratories and is affordable since it could be performed in one multiplex PCR reaction resulting in a reduced cost, less than 2 % of the price of an adult female. This test will allow parentage determination for different aims, particularly for pedigree establishment, laying the foundation for ongoing selection and genetic improvement, in the absence of systematic individual identification and rigorous recordings in the *Camelus dromedarius* herds.

#### 5. References

- Al-Atiyat R (2015).** The power of 28 microsatellite markers for parentage testing in sheep. *Electronic Journal of Biotechnology*. Volume 18, 2:116-121.
- Almathen F, Charruau P, Mohandesan E, Mwacharo JM, Orozco-terWengel P, Pitt D, Abdussamad AM, Uerpmann H, Uerpmann M, De Cupere B, Magee P, Alnaqeeb M. A, Salim B, Raziq A, Dessie T, Abdelhadi O. M, Banabazi M. H, Al-Ekna M, Walzer C, Faye B, Hofreiter M, Peters J, Hanotte O, Burger PA (2016).** Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. *Proceedings of the National Academy of Sciences*, 113 (24). pp. 6707-6712. ISSN 1091-6490.
- Da Silva EC, McManus CM, Guimarães MP, Gouveia AM, Facó O, Pimentel DM, Caetano AR and Paiva SR (2014).** Validation of a microsatellite panel for parentage testing of locally adapted and commercial goats in Brazil. *Genet. Mol. Biol.* Volume 37. 1:1415-4757.
- Douris R (1922).** Application medico-legal des groupes sanguins humains. *Discussion de la patemite*. *Bull. d. Sc.; Pharmacol.* 30: 90-107.
- Evdotchenko D, Han Y and Preuss S (2003).** New polymorphic microsatellite loci for different camel species. *Molecular Ecology Notes* 3:431-434.
- Heyen DW, Beaver JE and Da Y (1997).** Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semiautomated parentage testing. *Anim Genet* 28:21-7.
- Mariasegaram M, Pullenayegum S, Jahabar M, Shah RS, Penedo MC, Werery U and Sasse J (2000).** Isolation and characterization of eight microsatellite markers in *Camelus dromedarius* and cross-species amplification in *C. bactrianus* and *Lama pacos*. *Anim Genet* 33(5):385-387.
- Marshall TC, Slate J, Kruuk LEB and Pemberton JM (1998).** Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639-655.
- Miller SA, Dykes DD and Polesky HF (1988).** A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16 n°3.
- Nouairia G, Kdidi S, Ben Salah R, Hammadi M, Khorchani T and Yahyaoui MH (2015).** Assessing genetic diversity of three Tunisian dromedary camel (*Camelus dromedarius*) sub populations using microsatellite markers. *Emir J Food Agric* 27(4):362-366.
- Lang K, Wang Y and Plante Y (1996).** Fifteen polymorphic dinucleotide microsatellites in llamas and alpacas. *Anim Genet* 27:285-294.
- Luikart G, Biju-Duval MP and Ertugrul O (1999).** Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). *Animal Genetics*. 30:431-8.
- OEP (2016).** Rapport du plan national de développement du secteur camelin. *Officie de l'élevage et de pâturage*. Ministry of Agriculture and hydraulic resources and fishery of Tunisia.
- Ould Ahmed M, Ben Salem F, Bedhif S and Djemali M (2010).** Analyse moléculaire de la diversité génétique des dromadaires (*Camelus dromedarius*) en Tunisie. *Biotechnol Agron Soc Environ* 14(3):399-408.
- Penedo M, Caetano A, and Cordova K (1999).** Eight microsatellite markers for South American camelids. *Anim Genet* 30:161-168.
- Rendel J and Gahne B (1961).** Parentage tests in cattle using erythrocyte antigens and serum transferrins. *Anim Sc* 3:307-314.
- Sambrook J. and Russel D. (2001). *Purification of Nucleic Acids by Extraction with Phenol:Chloroform. Commonly Used Techniques in Molecular Cloning*, 3<sup>rd</sup> edn. Cold Spring Harbor Protocols, USA. Volume 3, Appendix 8.
- Spencer PBS, Wilson KJ and Tinson A (2010).** Parentage testing of racing camels (*Camelus dromedarius*) using microsatellite DNA typing. *Anim Gen*. doi:10.1111/j.1365-2052.2010.02044.x

- Stormont C (1967).** Contribution of Blood Typing to Dairy Science Progress. *J of Dairy Sc* 50(2):253-260.
- Wang J (2013).** Coancestry: a program for simulating, estimating and analyzing relatedness and inbreeding coefficients. *Mol Ecol Resour* 11(1):141-145.
- Weir BS (1979).** Inferences about linkage disequilibrium. *Biometrics* 35(1):235-254.
- Wright S (1978).** *Evolution and the Genetics of Population, Variability Within and Among Natural Populations.* The University of Chicago Press, Chicago.
- Yeh C, Boyle T, Rongcai Y, Ye Z and Xiyan JM (1999).** Microsoft Window-based Freeware for Population Genetic Analysis. Popgene. Quick user guide.