Medicine Papers, 2019, 5 (1): 3–8

A new application of Familias 3.0 Software in Veterinary Forensic Sciences

STEFANO REALE[1*], LORIANA GRAZIA VANESSA MASCALI[1,2], IGNAZIO SAMMARCO[1], TIZIANA LUPO[1], PIERLUIGI FERINA[1] & FABRIZIO VITALE[1]

[1]Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri", Via G. Marinuzzi 3, 90129 Palermo, Italy
[2]Indipendent Researcher Catania, Italy
[*]Corresponding author: stefano.reale@izssicilia.it

Summary

Familias software is commonly used in forensic laboratories to calculate probability of parentage between two or more individuals in cases of doubt relationships. Identifying family relationships through Familias software is well established in humans, but it is a hard task in animals due to high levels of inbreeding, lack of data about microsatellite frequency and other factors. In this study, we used Familias software for testing parentage relationships in a group of cattle through their microsatellite marker profiles. A standardized microsatellite panel recommended by the International Society for Animal Genetics (ISAG) was employed in conjunction with the allele frequencies database previously developed in a study involving 500 Sicilian cattle (Cosenza et al., 2015). Based on this bovine matrix and the known microsatellite profiles of the samples, the software permitted to calculate the Likelihood Ratios (LR) between different combination of parentage assignments and different pedigree hypotheses. The use of the software was validated and we report two of the most common veterinary forensic cases. The impact of this software on the quality of the analysis is critical in solving legal inquiry and animal forensic cases as much as the wrong calculation and data interpretation can invalidate the DNA profiling. Currently the described approach is useful for all the illegal inquiries (hundreds of cases) that judges entrusted us to resolve. Biological tests are often the only way to resolve these inquiries. The application of genetic identity tests could be complementary to the conventional traceability system, based on traditional labelling (ear-tag, bolus or transponder).

Key Words

Familias 3; Veterinary forensic sciences traceability; probability; inbreeding; kinship.

Received 23.10.2018; accepted 15.02.2019; printed 28.02.2019

Introduction

As well as in humans, also in animals the use of short tandem repeats as DNA markers represent a successful support for the identification of animals and carcasses and for parentage control. It's becoming also more and more important to trace the animals from the farmer to the slaughterhouse during the application of eradication programs for brucellosis, tuberculosis and leucosis. In veterinary forensic sciences we are often called during quality control programs for meat safety e traceability analysis. In the last years, microsatellite analysis has become the most consistent molecular approach in forensic sciences. Although the calculation of the probability of parentage relationships can be done by hand, in many cases it can be a really hard task, prone to human error (Brenner et al., 2004).

Many human forensic laboratories utilize a number of softwares as support in forensic tests, hence the possibility to identify dedicated softwares in veterinary field is now becoming necessary.

Since a child inherits half of his alleles from his father and half from his mother, different parent/child algorithms have been designed which search the DNA database for any profile that has one matching allele in common with the target profile at each locus.

This 'at least one allele per locus' matching approach was the first method used by forensic geneticists (Poulsen et al., 2013). Subsequent refinements led to include in the algorithms the population allele frequencies to rank the candidate lists by likelihood of relationship (Likelihood Ratio approach) (Buckleton & Triggs, 2006). Today, in forensic sciences the parentage or kinship between different individuals is calculated as likelihood-ratios (LR) between two different hypotheses (Dràbek, 2009). For instance, in a typical case the two competing hypotheses are:

- H1: ''Sample 1 directly matches (i.e. father, mother, sibling) with sample 2"

- H2: ''A random sample matches (i.e. father, mother, sibling) with sample 2''

A LR of 100,000 tell us that H1 hypothesis is 100,000 more likely than H2, assuming hence that Sample 1 and Sample 2 are truly related by a parentage relationship.

When we analyze animal kinship, for example in a herd, where inbreeding is extremely more common than in humans, the discrimination power of a determinate panel of microsatellites can be insufficient because of pour heterozygosity and low number of different alleles. The impact of a software, able to correct genotyping errors, taking into account inbreeding, null alleles or alleles dropout phenomena, is critical to correctly identify parentage relationships, and percentage of kinship. It is also easy to imagine how wrong calculation and subsequent data interpretation can lead to serious consequences and eventually to legal prosecutions.

The Familias 3 software is the gold standard tool when calculating likelihood ratios for different hypotheses of kinship given the genetic markers data of the samples (Gjertson et al., 2007), and is widely used by several forensic laboratories. Familias 3 may be used to compute probabilities and percentage of parentage in cases where DNA profiles of some samples are known, but their family relationship is in doubt (Egeland et al., 2000). The software is able to perform efficient computational analysis of family trees (or pedigree) with a moderate number of individuals (Huber, 1998). Given several alternative family trees for a group of samples and DNA measurements from some of these samples and taking into account a database of DNA observations in the relevant population (Thompson, 2000), the program may compute which relationship is most likely, and how much more likely it is than others. The software tool is able to handle complex cases, together with its ability to handle multiple pedigrees simultaneously. Familias includes different modules to perform parentage and identity testing. In particular: *Pedigree module* computes LR of different family trees, or pedigrees, given DNA observations of the studied samples (Egeland et al., 2000). The module computes which pedigree is most likely, and how much more likely it is than others; *Disaster Victim Identification* (DVI) module is used for identifying the biological family for a set of unidentified remains whose microsatellite profiles are known (Kling et al., 2014, 2017; Kling & Füredi, 2016); whitin DVI module the Blind Search function performs a survey about possible random kinship or unspecific relationships between the entire dataset.

This program has already been validated and is widely employed in humans (Dràbek, 2009) by many different forensic laboratories; its validation in veterinary forensic sciences is a new fact

and its application on breeded animals could contribute to the development of effective new avenues for parentage analysis, for traceability and certification of meats. The aim of this paper is to validate Familias 3 as a software able to identify, in a short time and with a low error rate, parentage and kinship in animals, determining the most probable relationships between two samples, given a certain set of genetic data.

In this paper, software validation requirements and tasks are defined, based on applying the available guidelines for the field of forensics (Buckleton, & Triggs, 2006) and software manual to the current needs of our field of investigation.

Material and Methods

Samples

A total of 11 cattle were analysed in the two main cases, 3 animals for case A and 8 for case B. In case B two blood samples were taken from each animal at two different times: at T0 in the herd and at T1 immediately before slaughtering, according to the brucellosis, tuberculosis, leucosis eradication programme for the traceability scope. The 16 samples (8 at T0 and 8 at T1) were numbered using the same ID for the two samples (from the farm or the slaughterhouse) from the same animal (i.e.from 1 to 8).

DNA extraction and fragment analysis

Although any kind of sample could be used for forensic purpose, we employed only whole blood samples in this study. All the times that the examined cases regard breeded living animals is preferable to

take blood from each identified animal. Genomic DNA was extracted and purified from all samples, using a commercial kit (PureLink Genomic DNA Mini Kit, Invitrogen, Carlsbad, CA, USA) according to the manufacturer protocol. DNA samples were stored at -20°C. DNA fragments were amplified in 11-plex PCRs using a certified commercial kit (StockMarks for Cattle Bovine Genotyping Kit, Applied Biosystems, Foster City, CA, USA) according to manufacturer instructions. The kit performed the amplification of 11 microsatellite loci among those recommended by ISAG - the International Society for Animal Genetics: TGLA 227, BM2113, TGL53, ETH10, SPS115, TGLA126, TGLA122, INRA023, ETH3, ETH225, BM1824 (Table 1). PCRs were carried out using a thermocycler (9700 Applied Biosystems, San Diego, CA, USA). Reactions were conducted in 15 ml final volume and the amplification program was optimized as follows: 10 min at 95°C, 31 cycles comprising 45 sec at 94°C, 45 sec at 61°C with a 50% speed ramp, 1 min at 72°C with 80% speed ramp. After a final polymerization step at 72°C for 1 hour followed by a hold at 25°C for 2 hours the test samples were maintained at 4°C. The PCR products were diluted with 135 µl of water before the injection on the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, San Diego, CA, USA). 1.5 µl of diluted amplified samples was mixed with 0.5 µl ROX size standard (Life Technologies) in 11 µl of deionized formamide running solution. The 3130 genetic analyzer was equipped with 35 cm capillary array and the filter set F was employed to reveal the FAM, JOE, NED dyes labelling the amplified alleles, and the size standard ROX dye. Genotypic profiles were read and analyzed using the GeneMapper ID v4.0 software (Applied Biosystems).

LOCUS	DYE	PEAK COLOURS	SIZE (bp)
TGLA227	FAM	Blue	64-115
BM2113	FAM	Blue	116-146
TGLA53	FAM	Blue	147-197
ETH10	FAM	Blue	198-234
SPS115	FAM	Blue	235-265
TGLA126	JOE	Green	104-131
TGLA122	JOE	Green	134-193
INRA23	JOE	Green	193-235
ETH3	NED	Yellow	90-135
ETH225	NED	Yellow	135-165
BM1824	NED	Yellow	170-218

Table 1. Microsatellites list with correspondent primer fluorescent dye. The size indicates the range in base pairs (bp) including all possible alleles per each locus.

Forensic cases

Two different forensic cases were reported in this work, investigating an overall of 11 cattle samples coming from three bovine farms of Palermo province, in a period ranging from October 2016 to February 2017. In the first case (Case A, property assignment), we investigated a suspected theft of a calf representing a common case of contentious between farmers. A breeder denounced the disappearance of an animal and its probable finding in a nearby farm. Because both the breeders declared the property of the animal and to own the relative mother, the judge ordered to us a genetic test of the alleged relatives (mother 1, mother 2, calf) (Fig. 1) to assess the true mother and hence the property of the calf. In this case we used the DVI - Blind Search module. In other occasions, using this approach we were able to add in our research a big amount of DNA samples, even two whole herds; however in this case, we used this module to found the real mother of the calf. The microsatellite profiles of Mother 1, Mother 2 and calf were analyzed and 50% of the alleles in common was flagged in searching relationship between cattles, giving us all the related couples with at least 50% of their alleles in common. Using a Bayesian approach the likelihoods were converted to posterior probabilities, taking into account allele frequencies of the reference population, which gave us important information about the possible kinship between the samples (parentage, direct matching, sister, etc.).

In the second case (Case B, meat safety e traceability), the judge asked the lab to check if some out of 8 cattle of a third farm (a different one from those envolved in case A) had been replaced before the slaughter and after having passed official controls. The traceability was resolved from a genetic point of view demonstrating the identity/non identity between the microsatellite patterns of two or more subjects. In this case 8 animals previously identified by official genetic controls were re-analyzed at the moment of slaughtering to demonstrate if the animals with the same ID code had the same microsatellite pattern too. A Blind search was employed as “direct match” to determine the genetic identity/not identity between two groups of samples taken from the same animals at two different times. Blind search functionality employs a “direct match” algorithm to determine if two

samples are taken from the same subject. The genetic identity can be useful to demonstrate that two samples belong to the same animal although they were taken in two different times. In forensic case B, 8 cattle were investigated at different moments to trace them and reveal an eventual substitution between animals before slaughter. In this case we first tried to identify if meat samples really came from the herd, by using the "Blind Search" module; when parent-child relationship was found, we then combined the "Pedigree module" to better confirm the result. Couples not identified by DVI analysis were considered as negative controls. A previously studied database of the microsatellite allele frequencies of Sicilian cattle (COSENZA ET AL., 2015) was included as part of the elaboration as reference data. For all possible comparisons Likelihood Ratio for the two competing hypotheses H1 and H2 was calculated by the formula:

$$LR = P(data|H1, \phi)/ P(data|H2, \phi)$$

where $P(data|H1, \phi)$ is the probability of hypothesis H1 given the microsatellite profile data and $P(data|H2, \phi)$ is the probability of hypothesis H2 given the same microsatellite profile data (WEIR, 1996).

RESULTS

Case A: property assignment

The Blind search module found a significant LR (196950.53) between the calf and the mother 1 coming from one of the farms (Fig. 2). The mother 2, coming from the other investigated farm, wasn't correlated to the calf because as suggested by the very low LR detected. The analysis gave to the court a very precise indication about the probable theft, revealing the true owner of the calf (Table 2).

Case B: meat safety e traceability

We were able to detect two substitutions of cattle before slaughtering. The first substitution involved the cattle 7 coming from the farm with the cattle 8 from the slaughterhouse, the second substitution involved the cattle 3 from the farm with the cattle 5 from slaughterhouse. In fact there was a correspondence between samples with different ID that should not be correlated, as demonstrated by LR values of 6.645e+020 and 8.248e+09, respectively. This result confirmed that two cattle were substituted before the arrival in the slaughterhouse. For all the other couples of samples involving the other animals, the microsatellite pattern were identical demonstrating as the animals at Slaughterhouse were the same of those at the farm (Table 3).

Property assignment	Relationship	LR
Calf / Mother 1	Parent/child	196950.53
Calf / Mother 2	Parent/child	n.r

Table 2. Results for the forensic case A. LR = Likelihood ratio of competing hypotheses; n.r. = not relevant.

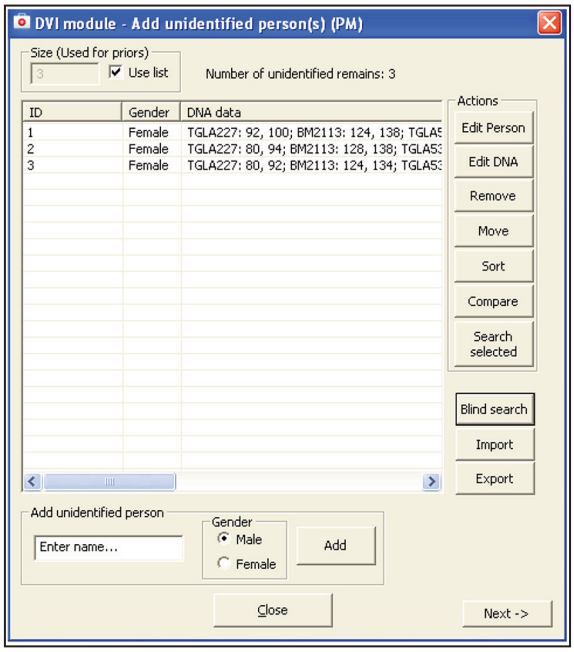

Figure 1. List of the three animals investigated for parentage testing with their DNA profiles.

CONCLUSIONS

In this paper, we have employed, for the first time, Familias 3 in veterinary forensic sciences. We demonstrated that the software born for human purpose, could be successfully used for parent/child and identity relationships in cattle. The software algorithms, based on Likelihood Ratio calculation between two different hypotheses, permit to have various corrections in case of population substructuring, high inbreeding rates, low heterozygosity and

Meat safety and traceability	**Relationship**	**LR**
Cattle 7 (Farm) Cattle 8 (Slaughterhouse)	Direct-match	6.6457608e+020
Cattle 3 (Farm) Cattle 5 (Slaughterhouse)	Direct-match	8.2481921e+009

Table 3: Results for the two forensic cases B. In the first column the identified matches.

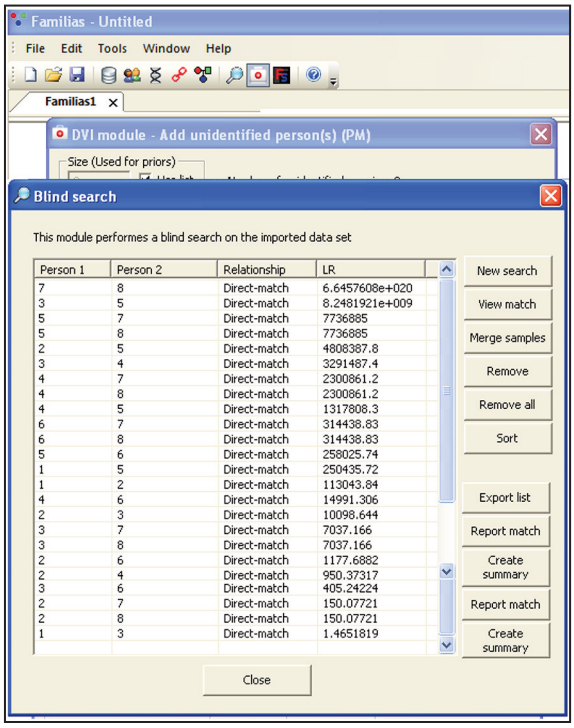

Figure 2. Blind search analysis for genetic identity/not identity determination between the samples (two per specimen, at two different times).

biodiversity in general. Using different modules and functionality included in Familias 3 software we applied a blind search based on microsatellite profiles of the investigated animals to solve two forensic cases involving some cattle from three different Sicilian herds in Palermo province. The different modules of the software were used to identify the true mother of a calf among two alleged mothers of different farms and to assess if some cattle have been replaced at the moment of the slaughtering with different individuals. Based on the allelic frequency matrix published in 2015 (COSENZA ET AL., 2015), indicating all the possible alleles for each locus, we were able to analyze and compare all our results by affirming or discriminating the grade of kinship or the genetic identity between the bovine samples. This could lead to a standardized methodology useful to obtain safe and indisputable data, especially if the analyses are involved in judgment processes. In addition, bringing the attention to this issue we would like to enhance the opportunity to improve this software for the specific employment in animal field. Permitting the implementation of the reference matrices, the algorithm can be considered as a dynamic structure adjustable on the basis of the cattle farmed species considered. Finally it permits a statistical approach at the cattle parentage and establish the basis to determinate the scoring in form of match probability to investigate both the parentage linkage and the identity through biological traces.

Focusing our attention to these issues, we developed a bioinformatic pipeline that will simplify forensic data analysis in lab and will reduce any kind of human error.

REFERENCES

BUCKLETON J. & TRIGGS C., 2006. The effect of linkage on the calculation of DNA match probabilities for siblings and half siblings. *Forensic Science International*, 160: 193–99. https://doi.org/10.1016/j.forsciint.2005.10.004

BRENNER C.H.. 2004. How to Solve Any Kinship Problem by Hand [Online], pp. 1–4. http://dna-view.com/index.html.

COSENZA M., REALE S., LUPO T., VITALE F. & CARACAPPA S., 2015. Allele frequencies of microsatellite loci for genetic characterization of a Sicilian bovine population. *Genetics and Molecular Research*, 14: 691–699. https://doi.org/10.4238/2015

DRÀBEK J., 2009. Validation of software for calculating the likelihood ratio for parentage and kinship. *Forensic Science International Genetics*, 3: 112–118. https://doi.org/10.1016/j.fsigen.2008.11.005

EGELAND T., MOSTAD P.F., MEVAG B. & STENERSEN M., 2000. Beyond traditional paternity and identification cases Selecting the most probable pedigree. *Forensic Science International*, 110: 47–59.

GJERTSON D.W., BRENNER C.H., BAUR M.P., CARRACEDO A., GUIDET F., LUQUE J.A., LESSIG R., MAYR W.R., PASCALI V.L., PRINZ M., SCHNEIDER P.M. & MORLING N., 2007. ISFG: Recommendations on biostatistics in paternity testing. *Forensic Science International Genetics*, 1: 223–231. https://doi.org/10.1016/j.fsigen.2007.06.006

Kling D. & Füredi S., 2016. The successful use of familial searching in six Hungarian high profile cases by applying a new module in Familias 3. *Forensic Science International Genetics*, 24: 24–32. https://doi.org/10.1016/j.fsigen.2016.05.012

Kling D., Tillmar A.O. & Egeland T., 2014. Familias 3 - Extensions and new functionality. *Forensic Science International Genetics*, 13: 121–127. https://doi.org/10.1016/j.fsigen.2014.07.004

Kling D. & Mostad P.F. & Egeland T., 2017. Manual for Familias 3. System 1, 78 pp.

Huber L., 1998. Qualification and validation of software and computer systems in laboratories. Part 2: qualification of vendors. *Accreditation and Quality Assurance*, 33: 2–5.

Poulsen L., Friis S.L., Hallenberg C., Simonsen B.T. & Morling N., 2014. A report of the 2009–2011 paternity and relationship testing workshops of the English Speaking Working Group of the International Society for Forensic Genetics. *Forensic Science International*, 9: e1–e2.

Thompson E.A., 2000. NSF-CBMS Regional Conference Series in Probability and Statistics Vol. 6. *Statistical Inference from Genetic Data on Pedigrees, Institute of Mathematical Statistics*, 169 pp.

Weir B.S., 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sunderland, MA: Sinauer Associates, 337 pp.