

Apr 22, 2020

Version 1

Analysis of genetic relatedness and paternity assignment in wild Guinea baboons (*Papio papio*) based on microsatellites V.1



DOI

<https://dx.doi.org/10.17504/protocols.io.bfdmji46>

Federica Dal Pesco¹, Franziska Trede¹, Dietmar Zinner¹, Fischer Julia¹

¹Cognitive Ethology Laboratory, German Primate Center, Göttingen



Federica Dal Pesco

German Primate Center

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.bfdmji46>

Protocol Citation: Federica Dal Pesco, Franziska Trede, Dietmar Zinner, Fischer Julia 2020. Analysis of genetic relatedness and paternity assignment in wild Guinea baboons (*Papio papio*) based on microsatellites. **protocols.io**

<https://dx.doi.org/10.17504/protocols.io.bfdmji46>

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited



Protocol status: Working

We use this protocol and it's working

Created: April 21, 2020

Last Modified: April 22, 2020

Protocol Integer ID: 35981

Keywords: Guinea baboons, *Papio papio*, microsatellite markers, genotyping, dyadic relatedness analysis, parentage analysis, paternity assignment in wild guinea, wild guinea baboon population, microsatellite marker, paternity assignment, genetic relatedness, microsatellite amplification, microsatellite, microsatellites this protocol, wild guinea, fecal sample, dna extraction,

Abstract

This protocol provides a standardized methodology for the analyses of genetic relatedness and paternity assignment in wild Guinea baboons (*Papio papio*) based on 24 microsatellite markers. It describes all methodological steps from sample collection and storage, DNA extraction, microsatellite amplification, and genotyping to a detailed explanation of the analyses regarding loci descriptive statistic, dyadic relatedness, and parentage analysis. This protocol was developed as a standardized method utilized by the Cognitive Ethology Laboratory (German Primate Center) to analyse fecal samples from the wild Guinea baboon population at the CRP Simenti field site, PNNK, Senegal.

1 **Sample collection and storage:**

We collected samples for genetic analysis non-invasively conducting fecal sampling of all identified individuals belonging to our study groups. Immediately after defecation approximately 5 g of fecal material was collected from the surface of the fecal bolus and placed inside a sterile tube containing 90% ethanol. Fecal samples were stored following a 2-step preservation protocol, with an initial phase of conservation in 90% ethanol followed by desiccation and storage with silica beads (Roeder et al., 2004). Samples were shipped to the German Primate Center in Germany every year and subsequently stored at -20°C .

2 **DNA extraction, microsatellite amplification, and genotyping:**

We extracted DNA from fecal samples using the First-DNA all tissue kit (Genial®). We characterized genetic variation in Guinea baboons by assessing the individual allele variation on 24 polymorphic autosomal microsatellite markers. The 24 markers were amplified in 5 different multiplex systems (mean number of loci per multiplex, 4.80 ± 1.10 SD) using the Multiplex PCR Kit (QIAGEN) and fluorescent-labelled primers in concentrations ranging from 0.05 to 0.5 μM . PCR cycling conditions included a hot start polymerase activation step at 95°C for 15min, followed by 40 cycles with denaturation at 94°C for 30s, primer annealing at 57°C for 40s, elongation at 72°C for 40s, and a single final elongation step at 72°C for 30min. Detailed information regarding loci identities, primer sequences and concentrations, and fluorescent dyes are reported in table 1. PCR products were separated and detected through capillary gel electrophoresis on an ABI 3130xL Genetic Analyzer (16 capillary sequencer, Applied Biosystems®, USA). Microsatellite allele sizes were evaluated in comparison with the GeneScan™ -400HD size standard using Gene Mapper 5 (Applied Biosystems®). Allele calling for each locus was repeated until 5 completely consistent calls were achieved. To avoid contamination, extraction, polymerase chain reaction (PCR) and fragment length analysis were performed in separate rooms and monitored with negative controls.

Table 1 - List of 24 microsatellite loci used in this study: multiplex-PCR, loci identities, primer sequences, fluorescence dye, and primer concentration

	Multi plex PCR	Locus ID	Primer 5'-3'	Fluo resc .	Primer Conc. (μM)
	M1	D6S26 4F	AGC TGA CTT TAC GCT GTT C	Fam	0.05
	57°C	D6S26 4R	TTT TCC ATG CCC TTC TAT CA	Fam	0.05
		D7S503 F	ATG ACT TGG AGT AAT GGG AG	Tam ra	0.15



		D7S503 R	GTC CCT GAA AAC CTT TAA TCA	Tam ra	0.15
		D12S37 5F	TTG TTG AGG GTC TTT CTC CA	Fam	0.09
		D12S37 5R	TCT TCT TAT TTG GAA AAG TAA C	Fam	0.09
		D3S176 6F	ACC ACA TGA GCC AAT TCT GT	Tam ra	0.05
		D3S176 6R	ACC CAA TTA TGG TGT TGT TAC	Tam ra	0.05
	M2	D14S30 6F	TCA GCT ACA TCC AAA TTA GGT	Tam ra	0.05
	57°C	D14S30 6R	TGA CAA AGA AAC TAA AAT GTC C	Tam ra	0.05
		D1S533 F	TAT CCC CCC CAA AAA TAT ATA	Fam	0.05
		D1S533 R	TTG CTA ACC AAA ATA ACA ATG GG	Fam	0.05
		D2S132 9F	TTG TAG AAC CCT CTC AAA TAT	Tam ra	0.5
		D2S132 9R	GAA ACT TCC ACC CTG GGT T	Tam ra	0.5
		D2S132 6F	AGA CAG TCA AGA ATA ACT GC	Hex	0.05
		D2S132 6R	CTG TGA CCC AAA AGC CGA	Hex	0.05
	M3	D10S61 1F	TAT ACA GGA AAC TGT GTA GTG	Tam ra	0.2
	57°C	D10S61 1R	CTA TAT TTA TGT GTG TGG ATG	Tam ra	0.2
		D8S110 6F	TTG TTT ACC CCT GCA CCA C	Hex	0.2
		D8S110 6R	TTC TCA GAA TTG CTC ATA GTG	Hex	0.2
		D17S79 1F	ATG TTC TCC AGT TAT TCC CC	Tam ra	0.5
		D17S79 1R	GCT GGT CCT TTG GAA GAG T	Tam ra	0.5
		D6S501 F	GCT GGA AAC TGA TAA GGG C	Hex	0.2



		D6S501 R	GCC ACC CTG GCT AAG TTA	Hex	0.2
		D17S12 90F	GAC AAC AGA GCA AGA CTG T	Fam	0.25
		D17S12 90R	AGA AGC AGT TAA ATG GCC AAA	Fam	0.25
		D6S311 F	ATG TCC TCA TTT GTG TTG TG	Tam ra	0.3
		D6S311 R	GAT TCA GAG CCC AGG AAG A	Tam ra	0.3
	M4	D5S145 7F	TAG GTT CTT GGC ATG TCT GT	Tam ra	0.2
	57°C	D5S145 7R	TGC TTG GCA TAC TTC AGG G	Tam ra	0.2
		D8S50 5F	CTA AAG TGA ACC CAA ACC TAA	Fam	0.15
		D8S50 5R	AGT GCT AAG TCC CAG ACC A	Fam	0.15
		D10S14 32F	CAG TGG ACA CCA AAC ACA AT	Tam ra	0.4
		D10S14 32R	TAG GTT ATC TAA ATA GTG GAT TT	Tam ra	0.4
		D5S82 0F	ATT GCA TGG CAA CTC TTC TC	Fam	0.3
		D5S82 0R	GTT CTT CAG AGA AAC AGA AC	Fam	0.3
		D3S176 8F	GGT TGC TGC CAA AGA TTA GA	Hex	0.15
		D3S176 8R	CAC TGT AAT TTG CTG TTG GAT	Hex	0.15
		D7S220 4F	TCA TGA CAA AAC AGA AAA TAA GT	Fam	0.4
		D7S220 4R	AGT AAA TGG AAT TGC TTG TTA C	Fam	0.4
	M5	D1S207 F	CAC TTC TCC TTG AAT CGC TT	Hex	0.1
	57°C	D1S207 R	GCA AGT CCT GTT CCA AGT C	Hex	0.1
		D4S24 3F	TCA GTC TCT CTT TCT CCT TG	Fam	0.15

		D4S24 3R	TAG GAG CCT GAG GTC CTG	Fam	0.15
		D1S548 F	GAA CTC ATT GGC AAA AGG AA	Hex	0.15
		D1S548 R	GCC TCT TTG TTG CAG TGA TT	Hex	0.15
		D21S14 42F	CTC CTC CCC ACT GCA GAT	Fam	0.5
		D21S14 42R	TCT CCA GAA TCA CAT GAG C	Fam	0.5

3 Genetic markers descriptive statistic:

We calculated descriptive statistics for all 24 markers, estimated F_{IS} , expected and observed heterozygosity, and tested for Hardy-Weinberg equilibrium (HWE) for all loci using the R package PopGenReport (Gruber and Adamack, 2015). We tested for the presence of null alleles using MICRO-CHECKER (Van Oosterhout et al., 2004). All loci that showed signs of null alleles and/or significant deviations from HWE were excluded from following analyses.

4 Genetic relatedness analysis:

We used the individual genotypes to estimate dyadic relatedness. As relatedness estimators are influenced by the allele-frequency distributions and the true relationships of the individuals included in the dataset (Blouin, 2003; Pew et al., 2015), we used the R package "related" (Pew et al., 2015; also see Wang, 2011) to choose the best relatedness estimator for each specific dataset and project. This package provides calculations and comparisons of five non-likelihood (Queller and Goodnight, 1989; Li et al., 1993; Ritland, 1996; Lynch and Ritland, 1999; Wang, 2002) and two likelihood estimators (Milligan, 2003; Wang, 2007). Based on the characteristics of the molecular markers of our dataset, we simulated 100 pairs of individuals each of known relatedness categories (parent-offspring, full-sibling, half-sibling, and unrelated). We then compared the observed and expected relatedness values for each of the seven relatedness estimators using Spearman's correlation and chose the estimator with the best performance and highest correlation coefficient. Depending on the type of estimator, the range of relatedness estimates and their interpretations varies. While the two maximum-likelihood estimators range from 0 to 1 (Wang 2007), the five non-likelihood estimators range from -1 to 1 and should be interpreted as a correlation as originally conceived by Wright (1921, 1922), where negative estimates indicate pairs that are less related than average (Wang, 2017).

5 Parentage analysis:

We assigned paternity using a likelihood approach with Cervus (Kalinowski et al., 2007). For all analyses we first ran a simulated parentage analysis with 10000 offspring followed

by the actual parentage analysis. Confidence levels of parent assignments were set to 95% ("strict" criterion; while the "relaxed criterion" was 80%). The following parameters were used in all calculations: proportion of potential sires sampled 0.95, proportion of typed loci 1.00, and proportion of mistyped loci 0.01. All females belonging to the study groups were included in this analysis and all potential transfers of mothers to another primary male's unit between time of conception and birth were examined. Identities of mothers were known from field observations, but we further checked all mother/offspring pairs with a maternity likelihood analysis (criteria for acceptance: identification as candidates with 0 mismatches). We ran a paternity analysis with a trio likelihood approach where the identity of the mother was known in order to determine the most likely father. For rare cases where the mother of the offspring was not known or not matching the field observations, the paternity analysis for this specific offspring was run not including mother identity. All adolescent (large juvenile and subadult) and adult males in our study were included as potential sires in this analysis. A male was considered to have sired an offspring when he was assigned as the most likely father, had 0 to maximum 1 mismatched alleles, and the confidence level for the assignments was more than 95%, according to the "strict" criterion.

6 References:

- Blouin, M.S., 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution*. 18, 503–511.
- Gruber, B., Adamack, A.T., 2015. landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers. *Molecular Ecology Resources*. 15, 1172–1178.
- Kalinowski, S.T., Taper, M.L., Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*. 16, 1099–1106.
- Li, C.C., Weeks, D.E., Chakravarti, A., 1993. Similarity of DNA fingerprints due to chance and relatedness. *Human Heredity*. 43, 45–52.
- Lynch, M., Ritland, K., 1999. Estimation of pairwise relatedness with molecular markers. *Genetics*. 152, 1753–1766.
- Milligan, B.G., 2003. Maximum-Likelihood Estimation of Relatedness. *Genetics*. 163, 1153–1167.
- Pew, J., Muir, P.H., Wang, J., Frasier, T.R., 2015. related: An R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*. 15, 557–561.
- Queller, D.C., Goodnight, K.F., 1989. Estimating Relatedness Using Genetic Markers. *Evolution*. 43, 258–275.
- Ritland, K., 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research*. 67, 175–185.
- Roeder, A.D., Archer, F.I., Poinar, H.N., Morin, P.A., 2004. A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes*.

4, 761–764.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*. 4, 535–538.

Wang, J., 2002. An estimator for pairwise relatedness using molecular markers. *Genetics*. 160, 1203–1215.

Wang, J., 2007. Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research*. 89, 135–153.

Wang, J., 2011. Coancestry: A program for simulating, estimating and analysing relatedness and in- breeding coefficients. *Molecular Ecology Resources*. 11, 141–145.

Wang, J., 2017. Estimating pairwise relatedness in a small sample of individuals. *Heredity*. 119, 302–313.

Wright, S., 1921. Systems of mating. *Genetics*. 6, 111–178.

Wright, S., 1922. Coefficients of inbreeding and relationship. *The American Naturalist*. 56, 330–338.