A REPORT ON PATERNITY TESTING IN BONOBOs (PAN PANISCUS): DIFFICULTY OF OBTAINING RELIABLE GENOTYPES BY ANALYSIS OF CA/GT REPETITIVE SEQUENCES

ABSTRACT: The paper shows that nuclear DNA extracted from plucked hairs can be used for the amplification of dinucleotide short tandem repeats (microsatellites). Five polymorphic microsatellite loci with CA/GT repeat units for paternity testing in a captive group of bonobos (Pan paniscus) were investigated. All PCR reactions performed yielded positive products at the first attempt. DNA extracted from hair is present in a small amount and is usually degraded. Therefore PCR products can often be easily obtained but the genotypes can be unreliable. Also in case of the present analysis, stutter bands (false alleles) were generated, making the assignment of true alleles difficult. It is concluded that these false alleles were results of reaction conditions during the amplification process. Methods used to determine the genetic relationships as well as possible consequences of the project for breeding management of the bonobo group are presented and discussed.

KEY WORDS: Pan paniscus – Hair samples – PCR – Microsatellite genotyping – DNA slippage

INTRODUCTION

Pan paniscus (Schwarz 1929), commonly known as the bonobo and previously often called pygmy chimpanzee, is one of only two existing species of the genus Pan. Bonobos are endemic to the Democratic Republic of Congo (formerly Zaire), south of the Congo River which forms a natural barrier separating them from their widespread congeners, common chimpanzees (Pan troglodytes). Because of continued deforestation of their habitat due to commercial lodging and illegal hunting maintained by the demand for bushmeat, bonobos are endangered today, their population decline continues and they inhabit only residues of their former ranges (Reinartz et al. 2000). Population size estimates for bonobos vary, ranging from as few as 5,000 to a possible 50,000 individuals (Dupain, Van Elsacker 2001).

Therefore, it is vital that genetic relationships are established in order to optimise outbreeding and heterozygosity in the captive population. Bonobo groups usually include several males and females in a party (Kano 1992). Therefore captive groups are to be kept under as natural conditions as possible, these conditions should include multi-male groups, as prevalent in the wild. Consequently, genetic analysis is usually necessary to establish kin relationships for optimal genetic management because it is not known which male has sired an offspring.

The use of non-invasive collection methods of genetic material (most often hair and faecal samples) has rapidly increased during the past decade. The non-invasive nature of these techniques is very beneficial because traditional methods of collection of genetic samples are based on capturing and immobilisation of animals and blood sampling. A derived advantage of non-invasive sampling is a lower risk of health related complications for both
animal and researcher. It is important to mention here that the introduction of these methods has been closely related to the development of various ways of DNA extraction for subsequent amplification by polymerase chain reaction (PCR) from minute amounts of the genetic material (see e.g. Gerloff et al. 1995, Vigilant 1999).

Microsatellites, generally known as Short Tandem Repeats (STRs), are now widely used for individual identification and tests of relatedness, and have become markers of choice in this area of application (Balding 1999). STRs comprise short repetitive nucleotide sequence motifs (1–8 bp in length) that are exceptionally polymorphic, making them very suitable for paternity exclusion analysis. In such analysis potential non-parents can be excluded if they do not possess an allele that the offspring must have inherited from them. Microsatellites have been found in every organism investigated so far and they occur as interspersed elements more or less uniformly distributed throughout the genome. CA/GT is the most common tandem repeat sequence in mammals and its abundance makes it readily detectable by PCR (Miller, Yuan 1997).

Nevertheless, a major concern about the use of dinucleotide repeats is their high propensity to an intramolecular mechanism, so-called DNA slippage (slipped-strand mispairing error of DNA polymerase), sometimes making the assignment of true alleles difficult (Schlötterer 1998). This effect has been studied in vitro as well as in vivo conditions. DNA slippage is induced by folding of the template strand into one or more loops during replication. DNA polymerase does not register these loops and synthesizes a new strand one or more repeat units shorter. In other words, deletion(s) occur(s). When folding occurs in the newly synthesized strand, it is consequently one or more repeat units longer. In other words, insertion(s) occur(s) (Levinson, Gutman 1987, Schlötterer, Tautz 1992). According to extra-band-appearance of some PCR products after their electrophoretic separation, these variations are usually referred to as "stutter" or "shadow" bands (Murray et al. 1993).

When comparing dinucleotide repeats with tetranucleotide ones, it has been found that a higher mutability is present in shorter units. In studies of various microsatellite loci it has been found that the possibility of generating replication loops is directly proportional to the number of microsatellite repeat units. However, it has not been discovered yet what is the minimum number of microsatellite units making possible that the DNA slippage operates. It is believed that there is also impact of sequences surrounding microsatellite on production of variations. However, consensus has not been found yet about whether the presence of G/C-rich sequences increases variability or not.

Several other factors have been proposed to influence a unique variability of repetitive sequences. Specifically, it is chromosomal location of the microsatellite that influences its instability (e.g. telomeric regions). The other important factors include the rate of methylation, temperature, phase of the cell cycle or homogeneity of the microsatellite. Heterogenous repeat sequences are more resistant to deletions and insertions than those with uniform repetitive motif (Ellegren 2000, Schlötterer 2000). PCR reaction conditions also play a very important role. The main influencing factors are pH and concentrations of dNTPs and Mg2+ that can change mutation ability of the DNA polymerase (Hite et al. 1996).

In the present study, genetic tests were required to assess paternity of the offspring because there was more than one sexually mature male in the group. An additional complication arose because a probable mother had not accepted her infant. Her pregnancy had not been unequivocally established and so maternity had to be genetically confirmed (see below).

Below, the methods used to resolve the genetic relationships are presented and difficulties of reliable assignment of alleles by analysis of dinucleotide microsatellite loci are discussed. Possible consequences of the project for the breeding management of the bonobo group at Twycross Zoo are also proposed.

<table>
<thead>
<tr>
<th>Name</th>
<th>Acronym</th>
<th>Gender</th>
<th>Date of birth</th>
<th>Age-class</th>
<th>Place of birth</th>
<th>Parents (M × F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diatou</td>
<td>DT¹</td>
<td>F</td>
<td>21/10/1977</td>
<td>adult</td>
<td>Stuttgart Zoo</td>
<td>MA × CA</td>
</tr>
<tr>
<td>Kakowet II</td>
<td>KA¹</td>
<td>M</td>
<td>07/06/1980</td>
<td>adult</td>
<td>San Diego Zoo</td>
<td>KW × LI</td>
</tr>
<tr>
<td>Kichele</td>
<td>KC</td>
<td>F</td>
<td>19/04/1989</td>
<td>adult</td>
<td>Stuttgart Zoo</td>
<td>MA × DT</td>
</tr>
<tr>
<td>Jasongo</td>
<td>JS¹</td>
<td>M</td>
<td>02/08/1990</td>
<td>adult</td>
<td>Wuppertal Zoo</td>
<td>MT × LL</td>
</tr>
<tr>
<td>Banya</td>
<td>BY¹</td>
<td>F</td>
<td>01/02/1990</td>
<td>adult</td>
<td>Cologne Zoo</td>
<td>? × BN</td>
</tr>
<tr>
<td>Keke</td>
<td>KE¹</td>
<td>M</td>
<td>02/01/1994</td>
<td>adult</td>
<td>Twycross Zoo</td>
<td>? × DT</td>
</tr>
<tr>
<td>Yasa</td>
<td>YS¹</td>
<td>F</td>
<td>27/08/1997</td>
<td>juvenile</td>
<td>Twycross Zoo</td>
<td>? × DT</td>
</tr>
<tr>
<td>Kinshasa</td>
<td>KY¹</td>
<td>F</td>
<td>19/10/2001</td>
<td>infant</td>
<td>Twycross Zoo</td>
<td>? × ?</td>
</tr>
<tr>
<td>Banbo</td>
<td>BB¹</td>
<td>F</td>
<td>03/09/2002</td>
<td>infant</td>
<td>Twycross Zoo</td>
<td>? × BY</td>
</tr>
<tr>
<td>Luo</td>
<td>LU¹</td>
<td>M</td>
<td>01/12/2002</td>
<td>infant</td>
<td>Twycross Zoo</td>
<td>? × DT</td>
</tr>
</tbody>
</table>

¹Hair samples obtained for genetic analysis.
²Data from Leus, van Puijenbroeck 2001.
MATERIALS AND METHODS

Study group
At the time of the study, the only bonobo group in the UK (Twycross Zoo, Leicestershire) was comprised of 3 sexually mature males (>7 years of age), 3 reproductively active females, 1 juvenile female and 3 infants (2 females and 1 male; Table 1). Keepers found the dominant female, DT, carrying the first-born infant, female KY. However, there was doubt whether DT was the mother because she had regular anogenital swellings prior to the birth of this infant. The lowest-ranking female BY, however, appeared weak and had heavy and irregular bleeding and a very irregular swelling cycle for several weeks after the birth (for more details see Vervaecke et al. 2003). The two other infants, male LU and female BB, were born to females DT and BY, respectively.

Genetic samples
Hair samples were obtained from all 9 relevant animals of the Twycross group in 2002 (Table 1; female KC has not reproduced at time of the study therefore she was not included in the genetic analysis). 12–15 plucked hairs with visible follicles were taken from each individual and placed in an unused paper envelope. Care was taken to avoid contaminating the samples by using surgical gloves and changing them with every new individual sampled. To further avoid possible contamination by human DNA, the envelopes were closed by cello tape, not by saliva. All samples were kept desiccated at room temperature for several weeks until DNA was extracted. Hairs that were not reproduced at time of the study therefore she was not included in the genetic analysis).

DNA extraction
Individual hair root segments were treated by Chelex ionic bead resin suspension with accompanying proteinase K digestion (Walsh et al. 1991). Two extractions were performed per individual in order to produce adequate template DNA for PCRs. DNA was extracted from 6 hairs per individual. Approximately, 3 mm from the root ends of hairs were cut and placed in 200 µL of 5% Chelex suspension (Biorad). After addition of 10 µL of proteinase K (10 mg/mL, Sigma), the samples were incubated overnight at 56°C with constant mixing, and then heated to 98°C for 10 minutes. Chelex beads were then sedimented into centrifugation into 11,000 g for 2 minutes. Two ~100 µL aliquots of Chelex bead free supernatant were transferred into new tubes. One tube from each extraction was kept at 4°C or –20°C, and the second at –80°C. The products of the two extractions were pooled together to constitute a 1:1 extract. Concentration of the DNA, measured by spectrophotometry, proved to be sufficient (C = 53.8 – 172.3 ng/µL). 0.5–3 µL of the final extract was used as template in each PCR reaction.

PCR amplification
Primers originally isolated in the context of the Human Genome Project were used to amplify microsatellite loci with PCR. These primers have been found to yield amplification products in a number of primate species (Gerloff et al. 1995, Warren et al. 2000, Lathuillière et al. 2001). A total of 5 CA/GT repeat loci were investigated: D1S207, D2S141, D6S271, D16S402, D17S791 (Table 2). Multiple amplifications of the loci were carried out in a 25 µL reaction (50 µM each dNTP, 1.5 mM MgCl₂, GeneAmp® 10X PCR Buffer, 0.1 U AmpliTaq™ Gold DNA polymerase (Perkin Elmer), 1mg/mL acetylated BSA (Promega), 10µM each of the primers (Sigma), 2.5 µL of the extract). PCR conditions for all the loci were as in Gerloff et al. 1995, using a PTC-200™ thermocycler (MJ Research). Initial denaturation was at 95°C for 11 minutes, followed by 7 cycles (93°C for 30 s, 50°C for 60 s, 72°C for 90 s) and 30 further cycles (89°C for 30 s, 54°C for 60 s, 72°C for 60 s).

Microsatellite genotyping
PCR products were visualised on 2% agarose gels stained with 0.5 µg/mL ethidium bromide to detect positive PCRs

<table>
<thead>
<tr>
<th>Locus</th>
<th>Repeat motif</th>
<th>Primer sequence¹</th>
<th>Allele size range²</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1S207</td>
<td>(CA/GT)ₖ</td>
<td>F 5'-CACCCTCTTCTGTGAAATCAGTCCTC-3' R 5'-GCAAGTCTGTTTCAAGTCTC-3'</td>
<td>142–170 bp</td>
</tr>
<tr>
<td>D2S141</td>
<td>(CA/GT)ₖ</td>
<td>F 5'-ACTACTACTACTACACTACACTACACTAC-3' R 5'-TTCTCAAAACAGATACAGTGAACCTT-3'</td>
<td>152–178 bp</td>
</tr>
<tr>
<td>D6S271</td>
<td>(CA/GT)ₖ</td>
<td>F 5'-AACATGGGAAATGGCTTA-3' R 5'-TAGGATGTGGTGGTGTTAC-3'</td>
<td>166–208 bp</td>
</tr>
<tr>
<td>D16S402</td>
<td>(CA/GT)ₖ</td>
<td>F 5'-TTTGTGAACTGATGACCC-3' R 5'-ATTTAGTGGCCATGACCC-3'</td>
<td>161–187 bp</td>
</tr>
<tr>
<td>D17S791</td>
<td>(CA/GT)ₖ</td>
<td>F 5'-GCTGTCCTGGTGAAGAGTGT-3' R 5'-GCTGTCCTGGTGAAGAGTGT-3'</td>
<td>165–199 bp</td>
</tr>
</tbody>
</table>

¹Primers were derived from human flanking sequences (Gyapay et al. 1994).
F – forward primer, R – reverse primer.
²Allele size ranges shown for all the loci are for humans (Gyapay et al. 1994).
According to Gerloff et al. (1995), the sizes of the amplification products in bonobos lie between 135 and 202 bp.

203
and to determine the dilution factor for the subsequent steps (run in 0.5X TBE; 80V/1h). All positive amplification products were electrophoretically separated on 8% polyacrylamide sequencing gels and detected with silver staining method (run in 0.5X TBE; 1900V/4h).

Then sequences of the D1S207 locus for 3 selected individuals (BY, KY, KE) were obtained by following procedure: fresh amplification products from 3 independent PCRs per individual were pooled together and purified by QIAquick® PCR Purification Kit (Qiagen), inserted in a vector by TOPO TA Cloning® Kit (with pCR® 2.1-TOPO®, Invitrogen) and recombinant plasmids were transferred into E. coli host cells. LB plates (LB medium, nutrient agar, kanamycin, X-gal) with the bacteria were incubated overnight at 37°C and then bacterial colonies containing recombinant plasmid were identified (white/blue identification). Plasmid DNA was liberated from the E. coli cells and purified by NucleoSpin® Plasmid Kit (Macherey-Nagel). EcoRI (Fermentas) cut PCR products with a short sequence of the plasmid surrounding the cloning site. 10 clones per individual were visualised on 3% agarose gel stained with 0.5 µg/mL ethidium bromide (run in 0.5X TBE; 80V/1.5h). This analysis of 10 clones (of which at least 7 need to be positive) was required to attain 99.2% probability of detecting both the alleles for a heterozygous individual (1 – 0.57 × 100 = 99.2, see e.g. Taberlet et al. 1996). After electrophoresis, only clones which differed in banding pattern were sequenced on ABI Prism® 3100 Avant Genetic Analyser using Big Dye Terminator v3.1 Cycle Sequencing Kit.

To be able to control the effect of PCR on arising false alleles we used DNA from successfully sequenced clones as templates for re-amplification of the D1S207 locus with the same PCR reaction conditions (only the allele with the highest number of the repeat units from each of the 3 above mentioned individuals was selected). Then PCR products were cloned and sequenced again.
FIGURE 3a, 3b. Results of the sequencing of the D1S207 locus. Only 3 individuals were tested (BY, KY and KE). All the sequences were double read and both readings provided the same data (results shown from 5’ to 3’). Only different alleles are shown. Sequence of the forward primer is underlined. Nucleotide positions are numbered from the beginning of the forward primer.

a – Results of sequencing of the positive clones after one PCR. As expected from results of electrophoresis with numerous shadow bands, more than 2 different alleles were generated for each individual, therefore some of the alleles must be false. 1–3 – alleles with different number of the CA/GT repeat units (in grey). Note that the KY’s allele 1 has the same length as BY’s allele 2, and KY’s allele 3 has the same length as the potential sire’s (KE’s) allele 2. b – Results of sequencing of the positive clones after two PCRs. 1 – true alleles with the highest number of the CA/GT repeat units which served as templates for second PCRs; 2 – PCR-generated (false) alleles. Note that the false alleles are in all cases one repeat unit shorter than the true alleles.
RESULTS AND DISCUSSION

For each of 9 individuals there were 2 DNA extractions. These were put together and 3 independent PCRs were carried out per individual and locus in order to amplify both alleles of heterozygous individuals (see Taberlet et al. 1996). All 135 PCR reactions performed yielded positive products at the first attempt (Figure 1). This is consistent with the idea that DNA extracted from hairs using Chelex resin can be detectable even though it is present in only small amounts. However, the quantity of DNA extracted is much lower compared with when blood or tissue samples are used as DNA source. It has been demonstrated that using hairs as a source of DNA may lead to unreliable genotype identification, even though it appears to be successful (Taberlet et al. 1996, Goossens et al. 1998, Goossens et al. 2000). Specifically, some alleles may not amplify (allelic dropout) and a heterozygote individual may appear as a homozygote, and/or PCR-generated alleles (false alleles) may arise.

After separation of the PCR products on polyacrylamide gels and their silver staining, too many bands became visible (Figure 2). Because gel electrophoresis detects all PCR products, irrespective of whether they are microsatellites or not, it was needed to discriminate between the microsatellite alleles and potential non-specific PCR products. It was for this reason that cloning and sequencing procedures were included into the study. Ten randomly chosen single DNA clones (molecules) per individual at the D1S207 locus were analysed (only for 3 individuals).

The sequencing confirmed that the amplified fragments that had been detected on the gel did indeed contain the microsatellite locus with its repetitive motif. Thus, it may be concluded that stutter bands had been generated with each amplification.

Unfortunately, we obtained more different sequences (microsatellite alleles) than expected for all 3 tested individuals. Interestingly, KY shares one allele with its probable mother BY and another one with one of its potential sires, KE (Figure 3a). Confirmation of KE's paternity of KY would be in agreement with earlier finding about his highest selective mating success with female BY (Stevens, unpublished data). However, when it is not known which allele(s) is (are) true, it is impossible to determine whether some two alleles are identical because there is a genuine relationship between their carriers, or whether these alleles are only "identical by state".

It has been found that the origin of false alleles is caused by the DNA slippage, occurring during the first cycles of the PCR and being subsequently amplified (Schlötterer, Tautz 1992, Ginot et al. 1996, Taberlet et al. 1996, and others). It has also been indicated that the most common mutation of a microsatellite allele is either the gain or loss of one repeat unit. Nevertheless, it is not possible to predict whether the next mutation will generate a longer or shorter microsatellite allele (however, see Xu et al. 2000). A re-amplification experiment showed that in cases of all 3 tested individuals it was possible to detect both true and false alleles. Interestingly, the false alleles were in all cases one repeat unit shorter than the true alleles (Figure 3b). Thus it may be concluded that the false alleles generated in this study were artefacts of the PCR reaction conditions. Moreover, the present results are in agreement with an empiric finding that microsatellites can be rather shortened than prolonged after the PCR (Miller, Yuan 1997).

For the above mentioned difficulties we have decided to orient our future analysis to 8 polymorphic tetranucleotide microsatellite loci for which the problem of stutter banding has not been documented (see Roeder et al. 2006). Another possible way how to reduce the occurrence of stutter bands seems to be the use of a DNA polymerase with correction activity (proof-reading), such as Pfu, Pfx or Pwo polymerases. It has been discovered that these polymerases do not synthesize so efficaciously in comparison to Taq polymerase, but their slippage error rates are much lower (e.g. Ginot et al. 1996). Therefore, it may be suitable to apply some of them together with Taq polymerase (in 1:1 ratio) during the PCR.

The results of this study are intended to benefit management of the bonobo group. In the wild bonobo females disperse, whereas males are philopatric and stay in their native community. Perhaps it is possible to emulate this pattern in Twycross Zoo, for example, by exchanging the young females with another zoo. Alternatively, when this does not seem practical, a new group could be established consisting of the young females. Under these circumstances it is crucial to establish which of the males in the group has sired the young females so that consanguineous sire-daughter matings and any resulting negative effect of inbreeding and loss of genetic variability can be avoided.

ACKNOWLEDGEMENTS

We are grateful to the administration and bonobo keepers of the Twycross Zoo for co-operating with this study. M. Smith and H. Dixon collected the hair samples. P. Brown and P. Milham provided important assistance with the work. Our thanks go to J. Mácha, T. Tlapáková T. Pexa and M. Zachová for their skilled assistance in the genetic analyses. We are also grateful to V. Fučík for helpful discussions. We wish to thank one anonymous reviewer for his advice and comments on earlier drafts of the manuscript. R. M. was financially supported by grants from the Ministry of Education, Youth and Sports of the Czech Republic (grant no. J13/98113100003), NATO Science Fellowships Programme (contract no. 13/2001), EuroMISE centre (grant no. LN00B107), foundation "Nadání J., M. a Z. Hlávkových", and East Midlands Zoological Society.
REFERENCES


