2

Determination of Genealogical Relationships from Genetic Data: A Review of Methods and Applications

Phillip A. Morin
Tony L. Goldberg

The concept of kinship has been central to investigating the remarkably varied social structures of primates. Genealogical relationships between individuals are predicted, from the first principles of evolutionary theory, to be critical influences on the nature of social relationships. Sociobiological/socioecological theory in particular predicts that kinship should have primary importance for the cohesion of groups, dominance, inbreeding avoidance, and coalitional behavior (Hamilton 1964, Wrangham 1980, Trivers 1985, Silk 1987).

Determining kinship is therefore a major focus of many studies of primate sociality. Kinship information has proven indispensable to addressing questions relevant to the evolution of sociality, mate choice, breeding systems, social dominance, and kin selection (Ross 2001). Consequently, investigators have used many direct and indirect methods to try to determine kinship relationships in primate groups. Among these, the most powerful, and currently the most widely used, are the molecular genetic methods.

Molecular genetic methods for determining kinship vary in their accuracy and in the amount of effort, expertise, money, and error involved. This chapter reviews molecular genetic methods that are commonly used or potentially useful in studies of primate kinship. Methods are reviewed with respect to their relative costs and benefits in terms of effort, financial cost, expertise, or specialized equipment, as well as with respect to the limitations of the inferences that can be drawn from them. A short review of the published applications of molecular methods for determining genealogical relationships follows to put the use of these methods into historical as well as methodological perspective.

The following sections describe potential sources of genetic material and the various classes of genetic markers commonly used in studies of primate kinship. Because of the current popularity and accessibility of DNA-based methods, we will not consider genetic methods that make use of other molecules (RNA, proteins). The goal is to provide informa-

tion that will facilitate the informed choice of an appropriate marker or set of markers. Please refer to table 2.1 for definitions of terms used throughout this chapter.

Genetic Methods

Sources of DNA

Molecular methods that make use of DNA without subsequent amplification by PCR require large amounts of high-quality starting material. Such DNA must be extracted from blood or tissue. These sources yield microgram to milligram quantities of high molecular weight DNA, which is required for direct visualization of the digested or probed fragments. To preserve the quality of the DNA during transport to the laboratory, such samples are typically stored and shipped frozen.

By contrast, PCR allows amplification of specific DNA fragments from as little as one copy of the genomic DNA template. PCR therefore expands the possible sources of DNA significantly, eliminating the need to collect large amounts of blood or tissue. Furthermore, samples collected for PCR need not be frozen immediately. Some degradation of the starting genetic material is acceptable, especially when the PCR amplicon is short (less than approximately 300 bp), as would be the case for most microsatellite loci and SNPs. Sources of DNA that have become accessible to primate geneticists since the invention of PCR now include shed or plucked hair (Vigilant 1999, Morin & Woodruff 1992, Morin et al. 1993, Goossens et al. 1998, Higuchi et al. 1988), feces (Taberlet et al. 1997, Launhardt et al. 1998, Gerloff et al. 1999, Immel et al. 2000, Smith et al. 2000), food wadges (Takasaki & Takenaka 1991), and, in theory, any other source that would contain genetic material from the primate of interest (Morin & Woodruff 1996, Taberlet et al. 1999).

Genomic Components and Modes of Inheritance

Primates have two sets of chromosomes in the nuclear genome, one set from each parent, and a single circular chromosome in each mitochondrion, which are present in thousands of copies per cell.

The mitochondrial genome is passed to offspring in the cytoplasm of the egg and is thus inherited only maternally. Because it is haploid (one copy) and inherited from one sex, the effective population size for mitochondrial genes is one-fourth that of autosomal (nonsex chromosomes) nuclear genes, which are diploid (two copies) and inherited from both parents. Mitochondrial DNA, therefore, reflects only matrilineality. Because of this, and because of its lack of genetic recombination, mitochondrial DNA is also useful for phylogeographic and phylogenetic studies, particularly when the sex bias of dispersal is known.

Nuclear, autosomal loci are found in two copies per cell. Individuals can be either homozygous or heterozygous at any given locus. Examination of nuclear autosomal loci therefore allows for the detection of heterozygotes within individuals, as well as the examination of variation among individuals in populations. Biparental inheritance of the alleles at these loci means that, when polymorphisms are present, the source of the variant alleles can be traced from parent to offspring (figure 2.1).

Table 2.1

AFLP

Allelic dro

Allele-spec
  hybridiz

Allozyme

Amplicon
Amplificat
Autosome
Codomina
Control re

Diploid

DNA fing

Electroph

Exon

Genotype

Haploid

Hv1

Intron

Locus

Mendelia

Microsat

Minisate

markers.

R require
ı blood or
ar weight
nents. To
; are typi-

tle as one
s of DNA
rthermore,
he starting
n approxi-
Sources of
PCR now
t al. 1993,
ıardt et al.
ˀakasaki &
terial from

ach parent,
ıousands of

and is thus
ɔne sex, the
nal (nonsex
m both par-
ıis, and be-
ır phylogeo-
known.
:ither homo-
ɔci therefore
ımination of
at these loci
:an be traced

Table 2.1. Useful Definitions

AFLP	Amplified fragment length polymorphism, a method for generating random, noncodominant variable genotypes based on restriction enzyme digestion and selective amplification of DNA fragments.
Allelic dropout	Artifactual loss of an allele from a genotype because of random PCR amplification of only one of two alleles (usually due to very small amounts of template DNA).
Allele-specific probe hybridization	Single-stranded DNA will hybridize with (bind to) other fragments of single-strand DNA to form double-stranded DNA when the two fragments are complementary. Under stringent conditions, hybridization will only take place when the two fragments are perfectly matched; alleles differing by ≥ 1 nucleotide can be distinguished by presence or absence of hybridization to DNA probes for each allele sequence.
Allozyme	A variant of a protein that can be differentiated based on protein size, charge, structure, or function, as detected by starch gel electrophoresis and various staining, denaturing, or enzymatic assays.
Amplicon	DNA fragment amplified via polymerase chain reaction (see PCR).
Amplification	Polymerase chain reaction DNA replication (see PCR).
Autosome	Diploid, nonsex chromosomes of the genome.
Codominant	Indicates that both alleles of a heterozygous genotype can be detected.
Control region	The region of the mitochondrial genome known as the origin of replication, and which does not code for a protein; also known as d-loop.
Diploid	Having two copies (maternal and paternal) of a chromosome (e.g., in the case of nuclear autosomal DNA).
DNA fingerprinting	Historically refers to the use of multilocus minisatellite probes to produce a pattern of DNA fragments that is often highly variable among individuals. Also used to describe composite multilocus genotypes.
Electrophoresis	A process of separating molecules by size (as in the case of DNA) or charge (as in the case of proteins) by passing them through a solid gel matrix using an electric current.
Exon	A section of protein-coding DNA. A gene may be made up of one exon, or many exons separated by introns.
Genotype	The combination of alleles present in an individual for a given locus or set of loci.
Haploid	Having one copy of a chromosome (e.g., in the case of mitochondrial DNA).
Hv1	The first hypervariable portion of the mitochondrial control region, most often used for phylogenetic and genealogical studies in primates and many other species.
Intron	A sequence of non–protein-coding DNA that lies between sections of protein-coding DNA (exons) that make up a gene.
Locus	A portion of the genome defined by function, location, or DNA sequence.
Mendelian heritability	Pattern of inheritance of one-half of the genome from each parent. When used to describe codominant alleles of a locus, this refers to the observed inheritance of one allele from each parent.
Microsatellite	Tandemly repeated short DNA sequence motifs. Repeated elements are usually between 1 and 6 nucleotides in length, and are repeated perfectly or imperfectly between 5 and 30 times. Also known as simple sequence repeats (SSRs) or simple tandem repeats (STRs).
Minisatellite	Tandemly repeated DNA sequence motifs, typically between 5 and 50 nucleotides long. Minisatellites are often near the telomeric (end) regions of chromosomes. Also known as variable number tandem repeats (VNTRs).

Table 2.1. Continued

MtDNA	Abbreviation for mitochondrial DNA, the circular, maternally inherited genome of the mitochondria.
Multilocus "fingerprint" or genotype	Compiled alleles from multiple loci that make up the composite genotypes for the individual.
Multiplex genotyping	The ability to combine multiple loci in a single assay and to distinguish individual genotypes from each locus.
PCR	Polymerase chain reaction, an enzymatic process of replicating specific fragments of DNA in vitro using cycles of DNA denaturation, primer annealing, and DNA polymerization (copying) of the template DNA strand.
Polymorphic locus	Any locus for which two or more alleles are present in a population.
Primer	Short synthetic DNA sequence (oligonucleotide) that is complementary to a segment of genomic DNA and is used to initiate replication of the DNA template during PCR.
Probe	DNA segment (oligonucleotide or larger DNA fragment) used to hybridize to a mixed or genomic DNA sample to identify that locus in the sample. The probe is usually labeled with radioactive material or other chemical attachments to facilitate visualization of the hybridized fragment on a solid matrix.
RAPD	Random amplified polymorphic DNA, a process by which short, nonspecific primers are used to amplify unknown regions of genomic DNA and thus to create individual-specific banding patterns.
RFLP	Restriction fragment length polymorphism, fragments of DNA created by the enzymatic hydrolysis of DNA based on the enzymatic recognition of specific DNA sequences.
Single nucleotide polymorphism (SNP)	Variation, within a population, in the nucleotide (A, C, G, or T) at a particular position in a DNA sequence.

FACING PAGE

Figure 2.1. Examples of electrophoretic detection of genetic variation in a pedigree. Black and hatched bands represent maternal and paternal alleles, respectively. Dashed bands can be either maternal or paternal and provide no information in these scenarios for parental exclusion. (A) variation in size (DNA) or charge (protein) of genetic markers. (B) Variation in DNA sequence detected by RFLP analysis. Change in the DNA sequence results in creation of a restriction enzyme recognition site, so the DNA is cleaved into two fragments at the new site. The heterozygous individual exhibits the patterns for the uncut fragment and the two cut fragments. (C) Variation in DNA sequence (SNP) for two loci detected by SBE and electrophoresis. Slight differences in allele size occur sometimes because of the different properties of the various fluorescent dyes attached to the SBE products. L1 and L2 refer to two multiplexed loci.

A
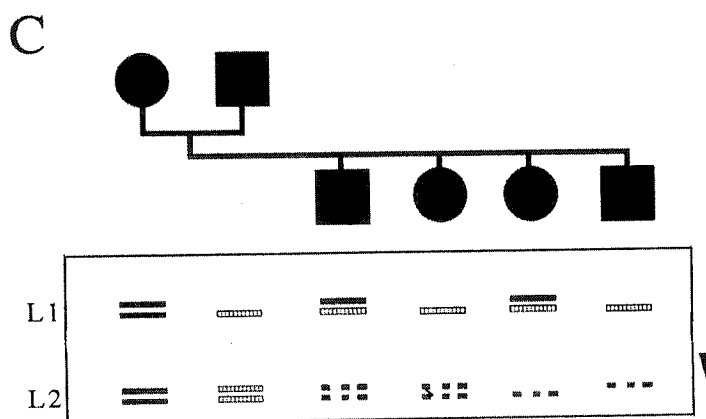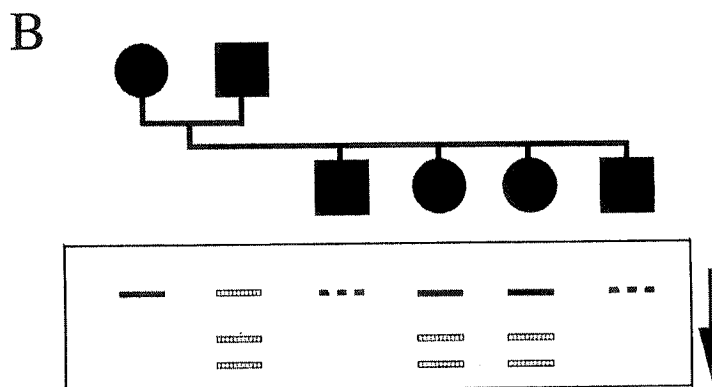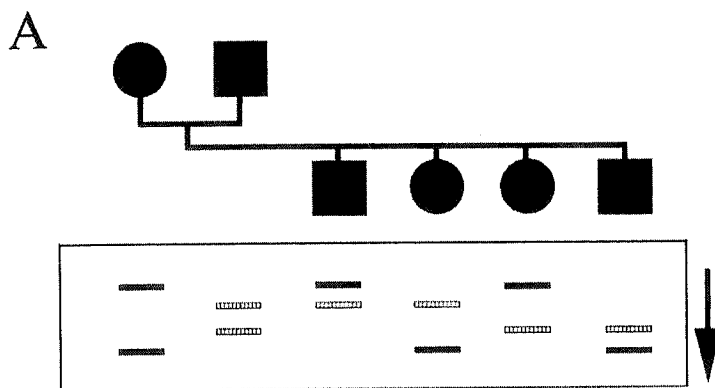
y inherited

site geno-

distinguish

ng specific
tion, primer
late DNA

pulation.
plementary
ication of

ed to hybrid-
ocus in the
erial or other
ridized frag-

hort, nonspe-
omic DNA

lA created
atic recogni-

r T) at a par-

B

C

igree. Black
d bands can
for parental
B) Variation
sults in cre-
fragments at
ragment and
cted by SBE
the different
l L2 refer to

L1

L2

In all primates, two of the nuclear chromosomes are sex chromosomes. The X and Y chromosomes are inherited in a sex-biased fashion, and are thus useful for studies of matrilineality and patrilineality, as well as for sex determination. Y chromosome genetic markers have been developed for humans and some nonhuman primate species (Underhill et al. 1997, Thomas et al. 1999, Stone et al. 2002). However, the process of selecting Y chromosome markers is difficult and time consuming, primarily because of the current paucity of species-specific Y chromosome DNA sequence data for primates other than humans. Nevertheless, Y chromosome markers could provide important information for male dispersal studies and for studies of kinship in males within and between groups (e.g., Thomas et al. 2000). X chromosome markers have not been reported in primate kinship studies to date and provide little information on relatedness that cannot be obtained from the more accessible autosomal markers or the more quickly evolving mitochondrial genome.

Sexing Primate Genetic Samples

The value of genetic information collected during primate kinship studies is often enhanced when the sex of individual samples can be determined. This is particularly true in noninvasive studies of unhabituated primates, when direct observation of individuals is impossible or unreliable.

Sexing of samples can easily be done for most primates using currently available techniques, or using minor adaptations of techniques developed for other mammalian species. Several methods have been described for PCR amplification of the amelogenin gene, for example, which is found on both the X and Y chromosomes, but which often differs in DNA sequence and length on the different chromosomes. A size polymorphism of 6 bp has been reported in humans (Sullivan et al. 1993), and is also present in other apes (Bradley et al. 2001). This assay has the advantage of producing a PCR product from one set of PCR primers for each sex chromosome, so that the assay is internally controlled for amplification. The disadvantage is that these primers have not been tested widely in primates, so it is not yet known whether they will amplify the targeted gene region, or whether the product will vary in size between the X and Y chromosome in other species. A similar assay involving size differences in the ZFX and ZFY genes has been reported for humans and a variety of New and Old World primates (Wilson & Erlandsson 1998). This assay is limited to high-quality DNA (e.g., from blood or tissue) because of the large size of the PCR amplification target.

An alternative method, which has been used on a wide range of mammals, makes use of PCR amplification of a segment of the SRY gene (Griffiths & Tiwari 1993). This assay has the advantage of being widely applicable among species. However, nonamplification of the PCR product is a nonspecific result, in that it can be due either to lack of a Y chromosome (female) or to failure of the PCR for other reasons. In this and other such analyses, one or more internal positive control PCRs should be performed simultaneously to validate the sexing results.

Genetic Marker Systems

Table 2.2 compares the most common or promising marker systems currently used to determine kinship relationships among primates. Genetic markers useful for inferring primate

Table 2.2. Comparison of Molecular Genetic Techniques Available to Study Kinship in Primates[a]

Table 2.2. Comparison of Molecular Genetic Techniques Available to Study Kinship in Primates[a]

	mtDNA Sequencing[b]	Multilocus Minisatellite	RAPD	AFLP[b]	Protein (Allozyme) Electrophoresis	RFLP (Nuclear Locus)	Microsatellites	SNPs[b] (SBE with Electrophoresis)
Development time[c]	1–4 weeks	1 month	1–2 weeks	2–4 weeks	2–4 weeks	1–4 months	2–6 months	2–6 months
Processing time[d]	2–4 weeks	2–4 weeks	1 week	1 week	2 days	2–4 weeks	2–4 days	2–4 days
Genetic variation	Medium (but variable)	High	Medium	Medium	Low (typically 2–4 alleles)	Low (typically 2–4 alleles)	High (typically 3–15 alleles)	Low (typically 2 alleles)
Risk of anomalous results	Low	Low–moderate	High (dominant alleles, sensitive PCR conditions)	Moderate–high (dominant alleles, sensitive PCR conditions)	Low	Low	Moderate (null alleles, allelic dropout)	Low
Ease of scoring	Moderate	Moderate–difficult	Difficult	Difficult	Easy–moderate	Easy	Easy–moderate	Easy
Relative cost[e]	High	Moderate	Low	Low–moderate	Low	Low	Moderate	Low–moderate

[a]Modified from Webster and Westneat (1998), with permission from Michael Webster and Brikhäuser Verlag.

[b]Marker types added to table by Webster and Westneat (1998), based on references and authors' experience (mtDNA, SNPs).

[c]Assumes lab is using the technique successfully on another species; if not, add 6–12 months minimum for all of the DNA techniques, 2–6 months for proteins to learn and develop procedures.

[d]For a set of 15–20 samples from collected tissue to bands ready to be scored (includes 1–4 days to isolate DNA for DNA techniques), for one probe, primer, or locus. For loci with low to moderate variability, multiple loci may be required, so time must be added accordingly. If noninvasive samples are used, sufficient replication also needs to be considered.

[e]Relative costs represent per-sample costs and do not include costs of marker development or specialized equipment.

kinship can be divided broadly into two classes, based on the type of laboratory method used. Single-locus methods examine individual loci separately and yield locus-specific genotypes. The major advantage of single-locus methods is analytical; alleles in single-locus systems are inherited in a Mendelian fashion and can be analyzed using powerful statistical methods derived from classical population genetic theory. Single-locus methods often suffer from relatively low individual locus variability, however, so that large numbers of loci are required to establish high confidence levels for unique individual or relationship identification.

Multilocus methods examine several loci simultaneously. Such methods generate patterns of DNA fragments, usually visualized as bands on a gel, that differ among individuals. The major advantage of multilocus methods is that they generate a great deal of information quickly. Their major disadvantage, however, is that the DNA fragments generated cannot be assigned to specific loci, and the data cannot be analyzed in a Mendelian framework. Interpretation of multilocus fingerprint patterns, especially beyond first-order relationships, is therefore difficult or impossible (Jeffreys 1987, Lynch 1988, Lander 1989, Jeffreys et al. 1991, Weatherhead & Montgomerie 1991, Smith et al. 1992; for a more complete review of these methods, see Martin et al. 1992, Smith & Wayne 1996, DeSalle & Schierwater 1998, Hoelzel 1998).

Genetic marker systems can also be divided into those that employ PCR and those that do not. This distinction is useful because the non–PCR-based methods require large amounts of high-quality starting DNA and are therefore of limited utility for studies of most wild primates. The non-PCR methods include, most notably, the classic "DNA fingerprinting" techniques so widely employed in early studies of paternity and reproductive success (see below). Nevertheless, these methods are being almost universally replaced by PCR-based techniques. We therefore concentrate on the PCR-based methods in this chapter.

Like molecular genetic methods in general, PCR-based methods can be divided into those that target specific, known loci, and those that generate random polymorphic DNA profiles from many loci simultaneously. The former include such methods as microsatellite and SNP genotyping, and are described in detail below. The latter include such methods as amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) analyses. The primary benefit of these multilocus PCR-based systems is that they do not require a priori knowledge of the DNA sequences of the target species. They have been widely used in plant and microbial genetics, but less often in animal studies. The primary problems with these systems are their unreliability with respect to reproducing consistent banding patterns, and the inherent interpretive limitation that the DNA fragments generated are not inherited in a Mendelian fashion. For these reasons, and because they have not been widely used in primate genetics, we will not consider these methods further; the interested reader can obtain more information from these review chapters: Caetano-Anolles (1998), Hoelzel and Green (1998), and Webster and Westneat (1998).

Methods considered further in this chapter are single-locus methods that are amenable to PCR. These methods have the widest potential utility for genetic studies of primate kinship, both in captivity and in the wild. The specific methods that we discuss include mitochondrial DNA sequencing, microsatellite genotyping, and single nucleotide polymorphism (SNP) genotyping. The first two methods are commonly used in primate kinship studies today, and the third method has great promise for the near future.

y method
ific geno-
gle-locus
statistical
ten suffer
f loci are
dentifica-

e patterns
luals. The
formation
ed cannot
amework,
tionships,
reys et al.
te review
hierwater

those that
e amounts
most wild
rprinting"
ccess (see
CR-based

vided into
phic DNA
rosatellite
methods as
phic DNA
s that they
They have
udies. The
ucing con-
fragments
cause they
ds further;
: Caetano-

amenable
rimate kin-
lude mito-
ymorphism
hip studies

The three methods have in common their ability to infer primate genotypes at specific loci. A genotype can mean anything from a DNA sequence to a set of allele sizes (microsatellites) or allele nucleotides (SNPs). Obtaining the genotypes represents the bulk of the work in relatedness studies and has been the goal of a correspondingly large variety of methods. For mtDNA sequencing and microsatellite genotyping, the most common methods are well established. For SNPs, the methodologies are evolving rapidly. Each marker type also has its own set of limitations and assumptions for application and analysis (for review, see Sunnucks 2000). The limitations described for each marker type below are especially pertinent to the estimation of genetic relatedness in nonhuman primates.

Mitochondrial DNA Sequencing

Mitochondrial DNA has several characteristics that recommend it strongly for some types of studies. From a practical standpoint, it is relatively easy to obtain from samples, even highly degraded ones, because of the high copy number in each cell. Its maternal inheritance, lack of recombination, and relatively high evolutionary rate also make it suitable for studies of phylogenetics and phylogeography, and for behavioral studies of matrilineality (e.g., Avise et al. 1987, Avise 1989, Morin, Moore, Chakraborty et al. 1994, Hashimoto et al. 1996, Goldberg & Wrangham 1997, Mitani et al. 2000, Pope 2000, Warren et al. 2001).

The mammalian mitochondrial genome contains 35 genes (13 protein-coding genes, 22 tRNAs) and the "control region" (also known as D-loop), or origin of replication, which does not code for a protein or RNA molecule and is highly variable. Different regions of the genome evolve at different rates and can thus be chosen to resolve "shallow" genealogical relationships (e.g., kinship) or "deep" relationships (e.g., systematics), as the situation warrants.

For intraspecific primate studies, the most commonly sequenced segment is the first hypervariable region (Hv1) of the control region (Hv1 is one of two hypervariable regions in the control region). Hv1 is popular as a target locus both because of its convenient size for PCR amplification (typically less than 450 bp), and because there are "universal" primers (Kocher et al. 1989) that work well to amplify it in a variety of species. As an example of the popularity of this segment, the HvrBase database (Burckhardt et al. 1999, www.hvrbase. org) for ape Hv1 sequences contained 9,309 human, 434 chimpanzee (*Pan troglodytes*), 4 bonobo (*P. paniscus*), 28 gorilla (*Gorilla gorilla*), and 3 orangutan (*Pongo pygmaeus*) Hv1 sequences as of September 2001. It should be noted that this database is not updated frequently and only contains unique sequences. Most published sequences are available through GenBank (http://www.ncbi.nlm.nih.gov/Entrez/).

Obtaining or Developing Mitochondrial Markers

The mitochondrial genome has been one of the most thoroughly studied segments of DNA in nonhuman species and has been sequenced in its entirety for more than 30 species, with the number growing on a monthly basis (e.g., Schmitz et al. 2000). Given the amount of information now available, it is relatively easy to find primers for taxa of interest in the literature or to design primers from published and aligned sequences available in public databases (e.g., AMmtDB: http://bighost.area.ba.cnr.it/mitochondriome, Lanave et al. 1999).

Furthermore, regions of mtDNA can be chosen based on the relative amount and patterns of variation desired for a particular study (Pesole et al. 1999).

The control region is the most variable portion of the mitochondrial genome and has been most widely used for investigations of relationships among individuals and groups below the species level. Primers for the entire control region (Kocher et al. 1989) or portions of it (e.g., Hv1 or Hv2) have been designed for various primate species. Nevertheless, sequencing of the control region from a variety of individuals of the target primate species may be required to ensure that species-specific primers are designed that avoid polymorphic sites and amplify the given segment from all (or the majority) of individuals.

Obtaining mitochondrial markers is therefore more straightforward than obtaining any other type of marker. The technical limitations are only those normally associated with DNA extraction, PCR amplification, and DNA sequencing.

Genotyping Methods for Mitochondrial DNA

Early methods of surveying variation in the mitochondrial genome primarily involved restriction fragment length polymorphism (RFLP) analysis. Direct DNA sequencing of PCR products (or cloned PCR products), however, has since become the method of choice (Hoelzel 1998). When little variation is present, some studies have screened for variants quickly using surrogate methods such as allele-specific probe hybridization (Morin et al. 1992) or other nucleotide screening methods (e.g., Amato et al. 1998, Dean & Milligan 1998).

Given the wide use of direct sequencing protocols, and companies or university core labs that perform sequencing efficiently and relatively inexpensively, the primary issues are (1) verifying that the sequence is truly mitochondrial and not a nuclear insert of a mitochondrial sequence (see below), and (2) choosing appropriate methods of sequence alignment and analysis.

Limitations and Assumptions of Mitochondrial DNA

The primary limitation of mtDNA in studies of genealogical relationship is that the entire mitochondrial genome is inherited intact (except for possible mutations) from the mother. As a result, mtDNA offers no information about paternity. Furthermore, the ability to use mtDNA for maternal lineage determination is dependent on the level of sequence variation in the population and its distribution; the fact that individuals in a group share mitochondrial haplotypes may or may not indicate that they are closely related. For example, some species have inherently low levels of variation (and therefore have shared haplotypes regardless of maternal relationship). Similarly, in some social systems, all females in a group may be maternally related. In these cases, mtDNA has little or no discriminatory power within groups. Nevertheless, patterns of mtDNA distribution on larger spatial scales may be very informative for inferring sex-biased dispersal and for making phylogeographic inferences.

Another problem with mtDNA sequence analysis is that portions of the mitochondrial genome have been incorporated into the nuclear genomes of most species. When the nuclear mitochondrial inserts (also called numts; Zischler et al. 1998, Zischler 2000, Bensasson et al. 2001) are amplified instead of the actual mitochondrial DNA, incorrect relationships can be inferred. This is particularly important for phylogenetic analyses, but can also lead to false inferences about individual relationships. Unfortunately, numts may be difficult or

impossible to distinguish from actual mitochondrial genes without prior information. Some phylogenetic analysis of the species of interest and its close relatives is usually required to determine whether mtDNA sequences are indeed from the mitochondrial genome. It has also been demonstrated that the amplification of numts may be more likely (in some species) from some DNA sources than from others (e.g., hair vs. blood; Greenwood & Pääbo 1999). Cloning of PCR products, followed by sequencing of multiple clones for each product, can sometimes reveal the presence of both the true mtDNA sequences and numts in a sample and therefore facilitate identification of the numts.

Finally, all studies to date that have used mtDNA to infer matrilineality have relied on DNA sequence data. Animals are assigned to different matrilines if they have different mitochondrial DNA sequences. Because typical mitochondrial sequences used in such studies are between 300 and 400 bp long, even error rates of less than 1% in DNA sequencing could lead to false exclusions of matrilineality. Methods of matrilineality exclusion that take such error into account are needed but have not been developed formally (Goldberg & Wrangham 1997).

Microsatellite Genotyping

The discovery in the 1980s of a class of highly variable nuclear markers, called simple sequence repeats (SSR), simple tandem repeats (STR), or microsatellites (Litt & Luty 1989, Tautz 1989, Weber & May 1989, Weber 1990), that are amenable to amplification by PCR, represented a major step forward in the analysis of individual genetic relationships in a variety of animal species (Queller et al. 1993). Microsatellite loci typically have 5 to 10 alleles varying in size by multiples of the repeat unit (e.g., two base pairs for a dinucleotide repeat, three for a trinucleotide repeat, and four for a tetranucleotide repeat). Microsatellites are abundant in the genomes of humans and many other species (Tautz 1989, Weber & May 1989; reviewed in Zane et al. 2001).

Because microsatellite alleles differ by size, they are amenable to detection and genotyping by polyacrylamide gel electrophoresis of the PCR products (figure 2.1). Initial detection methods involved radioactive tagging of the PCR products, but many methods for nonradioactive detection using fluorescent stains or modified nucleotides are now in use.

After initial studies demonstrated the utility of microsatellites in humans (Edwards et al. 1991, Moore et al. 1991, Schlötterer et al. 1991), their application to nonhuman primates followed quickly. Many loci were conserved between humans and the other primates, and could be amplified from noninvasive samples such as hair (Morin & Woodruff 1992, Morin et al. 1993). This was particularly important because of the desire of field primatologists to determine genetic relationships of their wild populations without disturbing the animals or reversing hard-won habituation.

Some of the first studies of wild primate populations using PCR amplification of microsatellites were in apes (Morin, Moore, Chakraborty et al. 1994, Morin, Moore, Wallis et al. 1994). These and future studies chose microsatellite loci from among those originally discovered in the human genome (Dib et al. 1996). Although this strategy worked well for apes and Old World monkeys (Morin & Woodruff 1992; Morin et al., 1997, 1998; Kayser et al. 1996; Coote & Bruford 1996; Wise et al. 1997), the discovery of markers de novo has been necessary for some species, especially New World monkeys and lemurs (Ellsworth & Hoelzer 1998, Jekielek & Strobeck 1999).

Obtaining or Developing Microsatellite Markers

Obtaining novel microsatellites for a species of interest often involves a laborious and complicated process of creating a genomic library (fragments of genomic DNA inserted into a bacterial or viral DNA vector) and enriching or screening it for the microsatellite repeats of interest (Zane et al. 2001 and references therein). For the majority of primate species, however, screening of previously characterized human markers is a more practical approach. Particularly for apes and Old World monkeys, this has resulted in the discovery of large numbers of highly variable microsatellites in a variety of species (e.g., Morin, Moore, Wallis et al. 1994, Coote & Bruford 1996, Ely et al. 1996, Kayser et al. 1996, Morin et al. 1997, Launhardt et al. 1998, Morin et al. 1998, Clifford et al. 1999, Goossens, Latour et al. 2000, Rogers et al. 2000, Smith et al. 2000, Zhang et al. 2001). The utility of this approach also extends to New World monkey species (Rogers et al. 1995, Ellsworth & Hoelzer 1998), although the success rate of finding variable loci may be lower and the incidence of null alleles (see below) higher. Given the recent increase in the number of companies able to make enriched microsatellite libraries to generate species-specific markers and the decrease in costs to have such libraries made, it may be more practical in some instances to employ a commercial service for this purpose.

Briefly, cross-species amplification involves the selection of a range of PCR conditions (usually covering approximately 10°C temperature range, and sometimes two to three magnesium concentrations from 1 to 3 mM), to try to amplify the target species DNA using human PCR primers. Sequence differences between human and target species DNA at the primer binding sites will reduce or prevent PCR amplification, however. By reducing the stringency of the PCR (lower temperature and/or higher magnesium concentration), reliable amplification of the homologous locus may be achieved despite the presence of some sequence differences. If the priming sites are conserved, one must then evaluate whether the locus itself is polymorphic or not in the target species.

Thousands of microsatellite loci are currently known in humans, and many of them can be purchased as unlabeled primer pairs for screening (e.g., Research Genetics, Huntsville, Alabama). For the majority of studies of nonhuman primates, it will be easier, faster, and more cost effective to screen human markers than to generate species-specific microsatellites using genomic libraries or other genome screening methods (Morin et al. 1998).

Fortunately, many primers have already been validated in nonhuman primate species. Although public databases are still not complete, the Molecular Ecology Notes database is a useful resource for finding published microsatellite markers (http://blackwellpublishing.com/Journals/men).

Genotyping Methods for Microsatellites

Microsatellite loci have the major advantage that alleles vary in size. Electrophoretic separation of DNA fragments is therefore perfectly suited to genotyping individuals at microsatellite loci. For fragments up to about 500 nucleotides, discriminating between fragments that differ in size by even a single nucleotide is possible using denaturing polyacrylamide gels, in conjunction with radioactive or fluorescent labels, or fluorescent or silver staining methods (David & Menotti-Raymond 1998, Schlötterer 1998). For this reason, PCR primers

that amplify microsatellite loci are generally designed to yield short amplification products (typically 100–300 bp).

The primary problems with microsatellite genotyping are accurate and reproducible sizing of the alleles, correct interpretation of the allele patterns for genotyping, and generation of accurate genotypes (not incorrectly determined due to contamination, null alleles, allelic dropout, or false alleles; see below). As stated above, genotype inaccuracy can be a major problem, especially when noninvasive samples are used. It is very important that researchers be aware of the possible problems, and that they take appropriate steps to avoid them or to compensate for them (e.g., Morin et al. 2001, Taberlet et al. 1996).

Limitations and Assumptions of Microsatellites

Microsatellites are currently the marker of choice for genetic studies of genealogical relatedness (for reviews, see Jarne & Lagoda 1996, Rosenbaum & Deinard 1998, Schlötterer 1998, Schlötterer & Pemberton 1998). Significant problems must be overcome, however, to ensure data integrity and quality. Among these are discriminatory power (see analysis section), null alleles (nonamplification of some alleles in the population), high mutation rates, and PCR artifacts (allelic stutter, dropout, and false alleles).

Null alleles (Callen et al. 1993) are difficult to detect and can cause incorrect assignment of homozygous genotypes to heterozygous individuals. Because PCR amplification depends on perfect or near perfect match of the PCR primers to the template DNA, nucleotide mismatches in the primer binding sites can cause PCR failure. When these mismatches are polymorphic in the populations, this can cause nonamplification of one allele in a genotype (failure of both alleles would be rarer, and may not be attributed to null alleles unless other reasons for PCR failure can be ruled out). Statistical tests for Hardy-Weinberg equilibrium (HWE) or verification of Mendelian heritability in known families can help to identify the presence of null alleles within a data set (Brookfield 1996).

Allelic dropout (Navidi et al. 1992, Taberlet et al. 1996, Gagneux, Boesch et al. 1997) can similarly produce apparent homozygotes. Allelic dropout results from the failure of PCR to amplify one or more microsatellite alleles at a locus, because of very low amounts of starting template. Studies that use noninvasive samples (e.g., feces, hair, food wadges) may suffer from both allelic dropout and false alleles (figure 2.2). Significant effort must be made to ensure accurate genotyping, including replicate PCRs of samples and/or quantitative evaluation of DNA content and integrity (Taberlet et al. 1996, Goossens et al. 1998, Goossens, Chikhi et al. 2000, Morin et al. 2001).

The high (and variable) mutation rates of microsatellite loci occasionally cause actual mutations (shifts from one allele size to another) between generations. In a study of kinship, this could result in false exclusion of an ancestor (e.g., a parent). Maximum likelihood analysis programs (e.g., Marshall et al. 1998, Goodnight & Queller 1999) employ methods of statistical analysis that correct for both mutation and user errors in the genotype assignment, so that reasonably robust assignments can be made regardless of some level of error in the data. High and variable mutation rates among loci can also be a problem for assessing population-level relationships (i.e., population structure), because models that estimate inter-populational distances from microsatellite data often assume particular mutation rates and mutational patterns (Balloux & Lugon-Moulin 2002).
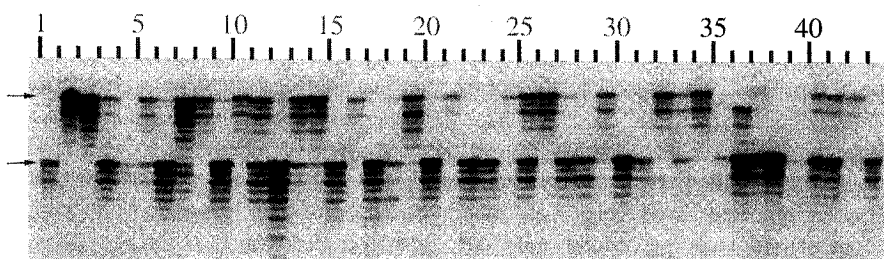
Figure 2.2. Replicate amplifications of the same sample for a single heterozygous locus (true alleles indicated with arrows), reprinted from Taberlet et al. (1996), with permission from Pierre Taberlet and Oxford University Press. This sample had a very low DNA concentration, so allelic dropout is the most common problem (e.g., lanes 1, 2, 3, 5), but false alleles are also evident (lanes 19, 37), and some allelic patterns are very difficult to interpret (e.g., lane 8). All of these problems can vary from locus to locus, and between samples, so choice of loci and quality of DNA are important issues to consider when a study is planned.

Single Nucleotide Polymorphism Genotyping

The type of polymorphism that occurs in the greatest abundance within primate genomes is the single nucleotide polymorphism. An SNP is simply a nucleotide position at which two (or, rarely, more) of the four possible bases (A, C, G, or T) occur within a population. SNPs are estimated to occur on average approximately every 1,000 bp in the human nuclear genome (Kruglyak 1997). SNPs have not been extensively exploited to date because most individual SNPs have only two alleles per site, and because of the limited technologies available for efficiently finding and genotyping many loci. Because of their low numbers of alleles per locus many more independent SNPs than microsatellites would need to be screened to achieve sufficient discrimination for kinship studies. Novel SNP detection is also labor intensive; de novo sequencing of many genome segments in many individuals is necessary. Finally, lack of size differences between alleles means that the SNPs cannot be detected simply by electrophoretic separation of PCR products.

The obvious way to generate SNP genotypes for animals would be to sequence large regions of nuclear DNA from many individuals within a population. The cost of such an endeavor would be high, however, even with today's technology. Fortunately, many methods for generating SNP genotypes have been developed, fueled by the search for genes involved in complex human disease (e.g., Marshall 1997). These methods are reviewed and updated almost monthly in journals.

Obtaining or Developing SNP Markers

SNPs can be found throughout the genome, but there is no way to predict SNP locations or to increase the likelihood of finding them (other than targeting genome segments that are less likely to be functionally constrained). Unlike microsatellites, which often show conserved regions of polymorphism in related species, SNPs may not be shared even between closely related species (Hacia et al. 1999). For this reason, and to obtain markers that are

40

ʒous locus
ɔermission
A concen-
, but false
ɔ interpret
amples, so
is planned.

ʒenomes is
which two
tion. SNPs
ɪuclear ge-
ause most
ɔhnologies
ɪumbers of
eed to be
etection is
ividuals is
cannot be

ence large
ɔf such an
ɪany meth-
for genes
·iewed and

ɔcations or
ts that are
show con-
n between
rs that are

not likely to be subject to strong selection, noncoding DNA is the most logical place to search for SNPs.

In an "unknown" mammalian genome, the method of choice for detecting novel SNPs has been to design PCR primers that bind to conserved regions of protein coding genes but amplify a product that includes an intron. These anchored primer pairs have been called comparative anchor tagged sequences (CATS; Lyons et al. 1997) or exon-primed intron-crossing (EPIC) loci (Palumbi & Baker 1994). Because the primers are designed to anneal to conserved regions, it is likely that they will amplify a homologous product in a wide variety of mammalian species. Recently, a subset of CATS (along with primers just for protein coding sequences) was used to create a nuclear gene phylogeny of the placental mammals, indicating that at least some of these loci will be widely useful (Murphy et al. 2001). At present, there are at least 200 such loci available to test (Venta et al. 1996, Lyons et al. 1997, Bruillette et al. 2000, Shubitowski et al. 2001), so it is likely that a sufficient number will amplify homologous segments in primates to allow rapid screening of sequences for novel SNPs.

Once PCR products are obtained for a given set of CATS, they can be sequenced from individual or pooled DNA samples to detect SNPs. The choice of method for SNP genotyping then dictates the steps for assay development and SNP verification (see below).

Genotyping Methods for SNPs

As mentioned, SNP genotyping is still relatively new and has yet to be applied in primate (or other species) genealogical studies. Critical issues for generating SNP genotypes for a relatively large number of loci and a population of individuals include capital equipment investment, assay component and reagent expenses, accuracy, and efficiency. Most methods are moving toward a "single tube" approach when possible, or at least minimizing the number of steps needed to generate the genotype data (for overviews of possible methods, see Kwok 2000, Gut 2001, Shi 2001, Syvanen 2001).

Since SNPs do not change the length of a DNA sequence, they cannot be detected by traditional electrophoresis. Some nucleotide changes can be detected by methods that combine electrophoretic migration of PCR products with changes in mobility due to denaturation (denaturing gradient gel electrophoresis, or DGGE) or secondary structure of the DNA (single strand conformation polymorphism, or SSCP) (Dean & Milligan 1998). These methods can be reasonably efficient and inexpensive, as they make use of equipment that is also used for sequencing or microsatellite genotyping and can reliably detect most polymorphisms. Not all polymorphisms, however, can be detected, and assays cannot be multiplexed, so each assay requires a separate lane on a gel or a separate capillary electrophoresis run.

A modification of the DNA sequencing chemistry called single base extension (SBE; also called minisequencing) makes use of modified (dideoxy) nucleotides with fluorescent labels to produce small oligonucleotides with fluorescent dyes that can be detected by electrophoresis or on microarrays. The method requires a PCR product large enough to encompass the SNP of interest and also provide a binding site for a single primer immediately "upstream" (5′) of the SNP nucleotide. Highly accurate polymerases that allow incorporation of a single dideoxynucleotide complementary to the SNP produce products labeled with a colored dye, which is different for each possible nucleotide (figure 2.3). These products can vary in size (depending on the oligonucleotide design) and color, so they can be multiplexed
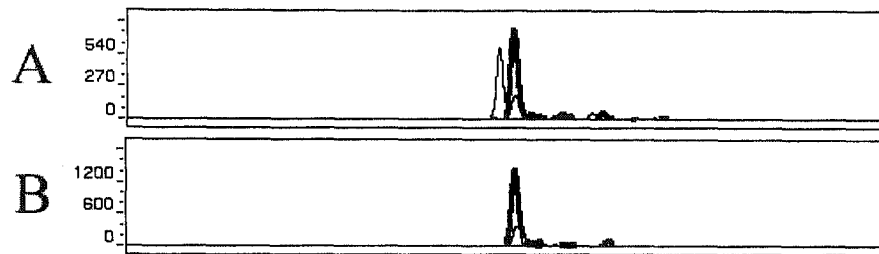
Figure 2.3. Single base extension genotype electropherograms for (A) heterozygous and (B) homozygous individuals for an SNP in the APOB gene. Small secondary peaks are thought to occur because of incomplete removal of primers after the initial PCR or the incomplete removal of dideoxynucleotides.

in electrophoresis for increased genotyping efficiency (e.g., SNaP-Shot SBE kit, Applied BioSystems, Foster City, California). SBE products can also be produced on microarrays (Syvänen & Landegren 1994, Pastinen et al. 1997), or produced in solution and resolved by hybridization to microarrays (Hirschhorn et al. 2000) or "liquid arrays" of microspheres (Chen et al. 2000). These methods require sophisticated equipment for fluorescent signal resolution on microarrays, but provide the opportunity for high levels of multiplexing and miniaturization for more efficient and inexpensive genotyping of many SNPs.

Finally, SNPs can be detected using probe hybridization methods, such as the 5' exo-nuclease assay (commonly called Taqman) or "molecular beacons." These methods use highly specific oligonucleotide probes that, under optimal conditions, hybridize only to exactly complementary sequences (Morin et al. 1999, Mhlanga & Malmberg 2001). The high specificity of these probes makes it possible to distinguish between DNA sequences that differ by even a single base pair within the region of probe hybridization (i.e., an SNP). Fluorescent dye systems can be used in conjunction with such probes so that their amplification can be tracked in real time and quantified during PCR amplification, or determined after amplification once systems have been optimized. The growing ease and accessibility of "real-time PCR" makes these approaches especially promising; genotypes can be generated in no more time than it takes to run a standard PCR, and loci can be multiplexed at a low level (two to four loci). These methods are highly reproducible and accurate, but assay reagents are relatively expensive, limiting their use for large numbers of assays, especially with small sample sizes.

Limitations and Assumptions of SNPs

To our knowledge, there have not yet been any demonstrations of the use of SNPs for genealogical relationship inference in primates, and it is not yet clear what the actual limitations will be. It is clear that the low information content of individual SNPs will mean that more SNPs than microsatellites will be needed to resolve kinship relationships. Theoretical evaluations have suggested that approximately 30 to 60 SNPs will be needed to match the power of a panel of 13 to 15 highly polymorphic microsatellite loci used in human forensics (Chakraborty et al. 1999, Krawczak 1999). Most SNP assay methods require amplification

of smaller DNA segments than microsatellites, so it is likely that amplification success from degraded DNA samples (e.g., noninvasive samples) will be higher, but this has yet to be shown. Null alleles will be a problem with SNPs as well (as with any PCR-based system); whether null alleles will prove to be less frequent for individual SNP loci than for microsatellites remains to be seen.

Data Analysis

Table 2.3 summarizes the applicability of various markers for studies of kinship. Below we discuss several applications in detail, including paternity exclusion, paternity inclusion, and determination of other kinship relationships.

Paternity Exclusion

The information content of a variable genetic marker is directly proportional to its level of variability (number of alleles) and inversely proportional to the variance in the frequencies of those alleles. For example, a locus that has five alleles of equal frequency is much more informative than a locus with two alleles, one of which is found at a frequency of 90%. This is because the probability that two individuals share an allele by chance alone is higher for the second locus (both individuals are highly likely to share the common allele).

Rare alleles can be very useful in tracing paternity or other relationships, since they are unique characters that are likely to be common to related individuals because of recent shared inheritance. To study many individuals, however, one must develop a suite of polymorphic markers that is likely to produce unique patterns in most or all cases (i.e., the probability of identity, Chakraborty et al. 1999). For paternity, this means assembling a set of markers that, on average, have a very low probability (usually less than 0.1%) of producing any particular pattern of alleles. In other words, two individuals chosen at random will share the haploid set of alleles (that half of the genome inherited from one parent, in the example of parent-offspring pairs) by chance less than one time in 1,000 comparisons (when $P < .1\%$). This is called the exclusion probability (E_H), and is expressed as the probability of excluding a random male who is not the father. Paternity can be considered excluded when E_H is very high (e.g., greater than 0.999).

For a given set of genetic loci, the exclusion probability can be calculated from the allele frequencies in the population (Chakravarti & Li 1983; reviewed in Morin & Woodruff 1992, Morin, Moore, Wallis et al. 1994). Estimates of E_H assume the population is in Hardy-Weinberg equilibrium. E_H is also strongly influenced by allele frequencies, so inaccurate estimates of the allele frequencies in the population will affect the estimates of exclusion probability for a given locus in the population. Nevertheless, if these conditions are satisfied, the individual probability of exclusion of an offspring can be calculated from its genotype and the frequency of alleles in the population (Chakraborty et al. 1988; reviewed in Morin & Woodruff 1992, Morin, Moore, Wallis et al. 1994).

As discussed above, different types of markers have different information content. Microsatellites have proven very useful for paternity studies because they are often highly variable, are present in most species, and can be genotyped without extremely sophisticated equipment. SNPs are much less variable, but are much more common in all genomes, and

Table 2.3. Applicability of Marker Types to Kinship Studies[a]

	mtDNA Sequencing	Multilocus Minisatellite	RAPD	AFLP	Protein (Allozyme) Electrophoresis	RFLP (Nuclear Locus)	Microsatellites	SNPs[b] (SBE with Electrophoresis)
Paternity (known mother)	N/A	Good–moderate	Good–moderate	Good–moderate	Moderate–low (need many loci)	Moderate–low (need many loci)	Good	Good–moderate (need many loci)
Siblings	Moderate	Moderate	Low	Low	Low	Low	Good–moderate	Good–moderate
Higher order relationship	Low	Moderate	Low	Low	Low	Moderate–low	Good–moderate	Low
Population assignment test	Low	Low	Low	Low	Moderate–low	Moderate–low	Good–moderate	Good–moderate
Population structure	Moderate	Low	Moderate	Moderate	Moderate–low	Moderate	Good–moderate	Good
Phylogeography	Good–moderate	Low	Low	Low	Good–moderate	Moderate	Low–moderate	Good–moderate

[a]For loci with few alleles/locus, the number of loci needed is higher. For allozymes and nuclear RFLPs, it is difficult to obtain a large enough number of loci, even though the resolving power is similar to that of SNPs, for which the possible number of loci is very high.

[b]The applicability is estimated by extrapolation from other codominant loci with few alleles, based on availability of large numbers of loci and projected ease of genotyping.

could potentially be genotyped more reliably and inexpensively (though the equipment may be more sophisticated and expensive for some methods).

The relative lack of information content in SNPs can be compensated for by adding more loci. In an analysis of the power of exclusion for a set of 13 standard human microsatellites used for forensics cases, Chakraborty et al. (1999) compared the power of those microsatellites to a hypothetical set of biallelic SNPs to see how many SNPs would be needed to obtain the same exclusion probability. Given a range of allele frequencies from 0.3 to 0.5 (for the less frequent allele), they concluded that 30 to 60 SNP loci would produce the same exclusion probability as these 13 microsatellites ($E_H > 0.99$). Similar conclusions were drawn by Krawczak (1999).

Paternity Inclusion

Exclusion should be used when the E_H is very high and all but one potential father can be excluded (Marshall et al. 1998). In many (but not all, e.g., Vigilant et al. 2001) field studies, this is not the case. When paternity exclusion is not possible in all cases, methods for paternity inclusion can be used.

Maximum likelihood (ML)–based methods for paternity inclusion allow the user to evaluate the data for the most likely father out of several candidates. These methods take into account, in addition to allele frequencies, the estimated genotyping error rates, the presence of null alleles, and the likely portion of potential sires not sampled. Maximum likelihood estimates calculate a statistical probability of paternity and evaluation of the next most likely candidate. This can be particularly important when potential sires are siblings, or when the information content of the markers, or the particular combination of alleles, leads to low resolving power between likely sires.

Because maximum likelihood methods depend on algorithms for fitting data to hypotheses and then comparing the fit, they require computer methods. Several software packages have been developed for paternity analysis, with a variety of options for data quality checking and statistical evaluation (e.g., Marshall et al. 1998, Goodnight & Queller 1999). These programs also typically calculate the average paternity exclusion probability for the loci screened on a given mother-offspring pair. The programs currently available have previously been reviewed in the context of statistical analysis of microsatellite data (Luikart & England 1999).

Analytical methods for estimating parentage are improving steadily. A new program called PAPA (Duchesne et al., 2002), for example, is available for parental allocation by likelihood methods when neither parent is known. This program also allows the user to simulate parental assignment given some knowledge of the allele frequencies in the population, and thus to determine whether the loci will provide sufficient power for parental allocation. This may be helpful for assessment of loci early in a study, thereby allowing researchers to find additional loci if needed.

Determining Other Classes of Relatedness

General patterns of relatedness of individuals in populations have been used to assess social structure, behavior, dispersal, and population structure (e.g., Morin, Moore, Chakraborty et al. 1994, Gagneux et al. 1999, Pope 2000, Constable et al. 2001, Vigilant et al. 2001).

Genetic markers have also been used to infer specific levels of relatedness, when actual genealogies are not known (reviewed in Queller et al. 1993). In practice, this involves determining the level of shared alleles in "unrelated" individuals of the population and in individuals of known relationship (e.g., parent-offspring, full siblings, half siblings), and then inferring the relationships of unknown pairs of individuals. This is a rapidly evolving field, and at least five different methods of inferring relationships from molecular data have been proposed (Queller & Goodnight 1989, Li et al. 1993, Blouin et al. 1996, Ritland 1996, Lynch & Ritland 1999; compared in Van De Casteele et al. 2001).

To date, however, such methods applied to primates have only been able to distinguish close (i.e., first-degree) relatives from other classes of relatives (Bruford & Altmann 1993, Altmann et al. 1996, Pope 1990, Gerloff et al. 1999, Vigilant et al. 2001). No genetic study of primate kinship has yet, in the absence of external data, been able to resolve relatedness on a finer scale (e.g., half versus full sibs, first versus second cousins). The molecular and analytical methods are progressing, however, such that the number of markers required for higher level relatedness inference will be reasonable, and the analytical methods refined for reasonable certainty of determining classes of relationships from samples within a group or population (Van De Casteele et al. 2001).

Applications of Molecular Genetic Methods to Determining Primate Kinship

The "holy grail" for genetic studies of primate kinship would be a complete account of genealogy for all animals in the social unit, extending back to its founding members. Such information would allow hypotheses about the influences of specific degrees of kinship on social behavior to be tested without genealogical error. Extension of such a genealogy across time and space would similarly inform studies of intraspecific differentiation, phylogeography, and population-level social organization.

In theory, a complete and accurate genealogy of living individuals could be reconstructed "blind" from genetic data alone. To date, however, this has never been done for any primate group. The number of variable, independently assorting genetic loci that would be required to achieve such a degree of precision would be impracticably large (Lynch 1988, Queller et al. 1993, Pemberton et al. 1999). Moreover, in all but captive and the most carefully studied wild primate groups, the required complete sampling of individuals would be difficult.

Fortunately, genealogy need not normally be reconstructed in the absence of prior information. Mother-offspring relationships can often be identified on the basis of behavioral observation alone (but see Smith et al. 1999). Behavioral observation can thus generate a preliminary matrilineal genealogy that can be resolved further using genetic data. Father-offspring relationships are less obvious. Especially in species with promiscuous mating systems, paternity is the one genealogical relationship for which behavioral inference should be considered unreliable a priori. Not surprisingly, therefore, the emphasis of DNA studies to date has been on resolving paternity.

Early primate paternity studies tended to focus on macaques (*Macaca* spp.) because of their ubiquity in captive settings. Also, largely as a result of work by Inoue and colleagues (e.g., Inoue et al. 1990, 1992), "DNA fingerprinting" methods (Southern blotting with minisatellite probes; Jeffreys et al. 1985, Jeffreys 1987) were available relatively early for macaques.

vhen actual
olves deter-
1 in individ-
1 then infer-
1g field, and
/e been pro-
1996, Lynch

) distinguish
tmann 1993,
;enetic study
∍ relatedness
olecular and
required for
ls refined for
n a group or

:e account of
embers. Such
of kinship on
ealogy across
phylogeogra-

reconstructed
or any primate
ld be required
88, Queller et
·efully studied
difficult.
of prior infor-
of behavioral
nus generate a
: data. Father-
us mating sys-
ference should
f DNA studies

ip.) because of
and colleagues
g with minisat-
· for macaques.

Many of the first such studies focused on the relationship between male dominance rank and reproductive success. The results of these studies were variable, sometimes showing strong positive associations between male rank and reproductive success and sometimes showing no association at all (see discussions in Turner et al. 1992, Bauers & Hearn 1994; also see Inoue & Takenaka 1993 and accompanying articles in that issue of *Primates*). Captivity seems insufficient for explaining this variability, since similar results were obtained in semicaptive (von Segesser et al. 1995, Bercovitch & Nurnberg 1997) and wild settings (de Ruiter et al. 1992, 1994; Keane et al. 1997). Positive associations were found most frequently when males maintained high-rank positions for sustained periods of time.

Traditional DNA fingerprinting technologies, as well as extensions of these using PCR and microsatellite loci, have since been applied to the elucidation of paternity in a variety of other species, including other Old World monkeys (Wickings & Dixson 1992, Bruford & Altmann 1993, Dixson et al. 1993, Altmann et al. 1996, Borries et al. 1999, Smith et al. 1999), New World monkeys (Dixson et al. 1992, Ellsworth & Hoelzer 1998, Pope 2000), and lemurs (Turner et al. 1992, Jekielek & Strobeck 1999, Fietz et al. 2000). Techniques for the noninvasive isolation of DNA from such sources as hair, feces, wadges, and chewed fruit have opened the door to studies of paternity in wild apes, which are not amenable to capture (Takasaki & Takenaka 1991, Morin & Woodruff 1992, Morin, Moore, Wallis et al. 1994, Gerloff et al. 1995, Field et al. 1998, Gerloff et al. 1999, Vigilant 1999, Bradley et al. 2000, Immel et al. 2000, Constable et al. 2001, Oka & Takenaka 2001, Vigilant et al. 2001).

Noninvasive paternity studies to date have been somewhat sobering, however. Morin, Moore, Wallis et al. (1994) investigated paternity in chimpanzees (*Pan troglodytes schweinfurthii*) in Gombe, Tanzania. Using PCR-amplified DNA from shed hair and a panel of eight microsatellite loci, they were able to assign paternity in 2 out of 10 cases and, in conjunction with behavioral data, to infer paternity in another 2 cases. Gagneux and colleagues (Gagneux, Woodruff et al. 1997, Gagneux et al. 1999) examined paternity in western chimpanzees (*P. t. verus*) from the Taï forest, Ivory Coast, for 13 infants of known maternity, using DNA from hair and chewed fruit and 11 PCR-amplified microsatellite loci. In seven cases, all community males could be excluded as fathers, indicating a surprising degree of extragroup paternity.

These results have been called into question, however, through analysis of additional and different loci, use of different analytical methods, and collection of new samples. Constable et al. (2001), using a likelihood-ratio approach, assigned paternity to three offspring in Gombe for which these fathers had previously been excluded. Vigilant et al. (2001), using new genetic data for an expanded study of the Taï chimps, were able to assign within-group paternity to four out of seven infants for which extragroup paternity was previously suspected. The phenomenon of allelic dropout, and subsequent false scoring of homozygosity at certain loci, apparently accounts for many of the initial erroneous paternity exclusions. The emerging consensus is that high replication and genotype verification measures to protect against contamination and false genotypes must be taken when primate kinship is inferred from DNA collected noninvasively (Taberlet et al. 1996, Morin et al. 2001).

Noninvasive paternity determinations in wild gorillas (Field et al. 1998), gibbons (*Hylobates muelleri*) (Oka & Takenaka 2001), and hanuman langurs (*Presbytis entellus*) (Borries et al. 1999) have failed thus far to document any significant discordance between social system and mating system. As longer term studies with larger sample sizes are conducted,

significant differences between social systems and mating systems may become evident. It will, therefore, be doubly important for these studies to use common genetic markers and genotyping methods that facilitate the comparison and combination of data sets, and to analyze these data using methods that allow for genotyping errors.

The use of genetic technology to elucidate kinship has not been limited to investigating paternity. For example, mitochondrial DNA (because of its unique mode of exclusively maternal inheritance) has been used to exclude and include matrilineality in lion-tailed macaques (*M. silenus*) (Morin & Ryder 1991), common chimpanzees (Goldberg & Wrangham 1997, Mitani et al. 2000), and bonobos (Hashimoto et al. 1996). These studies have shown that social preferences are largely independent of matrilineality in male chimpanzees and female bonobos respectively. Mitochondrial DNA can also be used to reconstruct matrilineality among females in female philopatric primates, when such relationships have not been documented behaviorally (Pope 2000).

To date, few primate studies have used genetic markers to examine kinship relationships other than paternity and matrilineality. Dixson et al. (1992), for example, confirmed the genetic identity of twins in wild common marmosets (*Callithrix jacchus jacchus*) using DNA fingerprinting. Morin, Moore, Chakraborty et al. (1994) and Vigilant et al. (2001) have used allele-sharing methods to examine levels of relatedness among groups of individuals within and between chimpanzee communities. De Ruiter and Geffen (1998) used blood protein markers to infer relatedness among long-tailed macaques (*M. fascicularis*) for which external pedigree data were available. No behavioral studies to date have used Y chromosome data to determine patrilineal relatedness, although this would be feasible.

Genetic methods are well suited to the reconstruction of all classes of kinship, including those that span temporal and spatial scales great enough that their study would more conventionally be considered phylogeography. With the advent of technologies that can examine large numbers of variable loci simultaneously, "blind" genealogy reconstruction may become possible in the future. In the interim, however, the best studies are those that test specific hypotheses about specific degrees of kinship, and that choose genetic markers appropriate to the question at hand.

Conclusion

Behavioral studies tell us a great deal about primate societies, but actual kinship relationships are often difficult or impossible to infer without genetic information. Within the last two decades, technical and analytical advances in genealogy and relatedness detection methods have made it possible for any organism or population to be studied, and for kinship to be resolved at most levels.

Nevertheless, the techniques available today do not have the resolving power to reconstruct kinship reliably in the absence of external information. Genetic markers for study must therefore be chosen wisely, with consideration of their technical limitations and their ability to resolve kinship at the level in question. The best studies are those that define a specific kinship-related hypothesis and choose a marker system suitable for testing it.

for generating the SNP genotype figure. Pierre Taberlet and Oxford University Press granted permission to reproduce the allelic dropout gel image, and Michael Webster and Birkhäuser Verlag granted permission to reproduce and modify table 2.2.

References

Altmann, J., Alberts, S. C., Haines, S. A., Dubach, J., Muruthi, P., Coote, T., Geffen, E., Cheesman, D. J., Mututua, R. S., Saiyalel, S. N., Wayne, R. K., Lacy, R. C., & Bruford, M. W. 1996. Behavior predicts genetic structure in a wild primate group. *Proc. Natl. Acad. Sci. USA*, 93, 5797–5801.

Amato, G., Gatesy, J., & Brazaitis, P. 1998. PCR assays of variable nucleotide sites for identification of conservation units: an example from *Caiman*. In: *Molecular Approaches to Ecology and Evolution* (Ed. by R. DeSalle & B. Schierwater), pp. 177–190. Boston: Birkhäuser.

Avise, J. C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution*, 43, 1192–1208.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.*, 18, 489–522.

Balloux, F. & Lugon-Moulin, N. 2002. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.*, 11, 155–165.

Bauers, K. A. & Hearn, J. P. 1994. Patterns of paternity in relation to male social rank in the stumptailed macaque, *Macaca arctoides*. *Behaviour*, 129, 149–176.

Bensasson, D., Zhang, D. X., Hartl, D. L., & Hewitt, G. M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.*, 16, 314–321.

Bercovitch, F. B. & Nurnberg, P. 1997. Genetic determination of paternity and variation in male reproductive success in two populations of rhesus macaques. *Electrophoresis*, 18, 1701–1705.

Blouin, M. S., Parsons, M., Lacaille, V., & Lotz, S. 1996. Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.*, 5, 393–401.

Borries, C., Launhardt, K., Epplen, C., Epplen, J. T., & Winkler, P. 1999. Males as infant protectors in Hanuman langurs (*Presbytis entellus*) living in multimale groups: defense pattern, paternity and sexual behaviour. *Behav. Ecol. Sociobiol.*, 46, 350–356.

Bradley, B. J., Boesch, C., & Vigilant, L. 2000. Identification and redesign of human microsatellite markers for genotyping wild chimpanzee (*Pan troglodytes verus*) and gorilla (*Gorilla gorilla gorilla*) DNA from faeces. *Conserv. Genet.*, 3, 289–292.

Bradley, B. J., Chambers, K. E., & Vigilant, L. 2001. Accurate DNA-based sex identification of apes using noninvasive samples. *Conserv. Genet.*, 2, 179–181.

Brookfield, J. F. Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.*, 5, 453–455.

Bruford, M. W. & Altmann, J. 1993. DNA fingerprinting and the problems of paternity determination in an inbred captive population of Guinea baboons (*Papio hamadryas papio*). *Primates*, 34, 403–411.

Bruillette, J., Andrew, J., & Venta, P. 2000. Estimate of nucleotide diversity in dogs with a pool-and-sequence method. *Mamm. Genome*, 11, 1079–1086.

Burckhardt, F., von Haeseler, A., & Meyer, S. 1999. HvrBase: compilation of mtDNA control region sequences for primates. *Nucleic Acids Res.*, 27, 138–142.

Caetano-Anolles, G. 1998. Arbitrary oligonucleotides: primers for amplification and direct identification of nucleic acids, genes and organisms. In: *Molecular Approaches to Ecol-*

ogy and Evolution (Ed. by R. DeSalle & B. Schierwater), pp. 107–123. Berlin: Birkhäuser Verlag.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C., & Sutherland, G. R. 1993. Incidence and origin of "null" alleles in the $(AC)_n$ microsatellite markers. *Am. J. Hum. Genet.*, 52, 922–927.

Chakraborty, R., Meagher, T. R., & Smouse, P. E. 1988. Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. *Genetics*, 118, 527–536.

Chakraborty, R., Stivers, D. N., Su, B., Zhong, Y., & Budowle, B. 1999. The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis*, 20, 1682–1696.

Chakravarti, A. & Li, C. C. 1983. The effect of linkage on paternity calculations. In: *Inclusion Probabilities in Parentage Testing* (Ed. by H. F. Polesky), pp. 411–420. Arlington, VA: American Association of Blood Banks.

Chen, J., Iannone, M. A., Li, M. S., Taylor, J. D., Rivers, P., Nelsen, A. J., Slentz-Kesler, K. A., Roses, A., & Weiner, M. P. 2000. A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. *Genome Res.*, 10, 549–557.

Clifford, S. L., Jeffrey, K., Bruford, M. W., & Wickings, E. J. 1999. Identification of polymorphic microsatellite loci in the gorilla (*Gorilla gorilla gorilla*) using human primers: application to noninvasively collected hair samples. *Mol. Ecol.*, 8, 1556–1558.

Constable, J., Ashley, M. V., Goodall, J., & Pusey, A. E. 2001. Noninvasive paternity assignment in the Gombe chimpanzees. *Mol. Ecol.*, 10, 1279–1300.

Coote, T. & Bruford, M. W. 1996. Human microsatellites applicable for analysis of genetic variation in apes and Old World monkeys. *J. Heredity*, 87, 406–410.

David, V. A. & Menotti-Raymond, M. 1998. Automated DNA detection with fluorescence-based technologies. In: *Molecular Genetic Analysis of Populations: A Practical Approach* (Ed. by A. R. Hoelzel), pp. 337–370. New York: Oxford University Press.

Dean, M. & Milligan, B. G. 1998. Detection of genetic variation by DNA conformational and denaturing gradient methods. In: *Molecular Genetic Analysis of Populations: A Practical Approach* (Ed. by A. R. Hoelzel), pp. 263–286. New York: Oxford University Press.

de Ruiter, J. R. & Geffen, E. 1998. Relatedness of matrilines, dispersing males and social groups in long-tailed macaques (*Macaca fascicularis*). *Proc. Royal Soc. Lond., B*, 265, 79–87.

de Ruiter, J. R., Scheffrahn, W., Trommelen, G. J. J. M., Uitterlinden, A. G., Martin, R. D., & van Hooff, J. A. R. A. M. 1992. Male social rank and reproductive success in wild longtailed macaques: paternity exclusions by blood protein analysis and DNA fingerprinting. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 175–191. Basel: Karger.

de Ruiter, J. R., Van Hooff, J. A. R. A. M., & Scheffrahn, W. 1994. Social and genetic aspects of paternity in wild long-tailed macaques (*Macaca fascicularis*). *Behaviour*, 129, 203–224.

DeSalle, R. & Schierwater, B. 1998. *Molecular Approaches to Ecology and Evolution.* Boston: Birkhäuser.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., & Weissenbach, J. 1996. A comprehensive genetic map of the human genome based on 5264 microsatellites. *Nature*, 380, 152–154.

Dixson, A. F., Anzenberger, G., Monterio Da Cruz, M. A. O., Patel, I., & Jeffreys, A. J.

1992. DNA fingerprinting of free-ranging groups of common marmosets (*Callithrix jacchus jacchus*) in NE Brazil. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 192–202. Basel: Karger.

Dixson, A. F., Bossi, T., & Wickings, E. J. 1993. Male dominance and genetically determined reproductive success in the mandrill (*Mandrillus sphinx*). *Primates*, 34, 525–532.

Duchesne, P., Godbout, M. H., & Bernatchez, L. 2002. PAPA (Package for the Analysis of Parental Allocation): a computer program for simulated and real parental allocation. *Mol Ecol Notes,* 2, 191.

Edwards, A., Civitello, A., Hammond, H. A., & Caskey, C. T. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.*, 49, 746–756.

Ellsworth, J. A. & Hoelzer, G. A. 1998. Characterization of microsatellite loci in a New World primate, the mantled howler monkey (*Alouatta palliata*). *Mol. Ecol.*, 7, 657–658.

Ely, J., Campbell, J. L., Gonzalez, D. L., & Stone, W. H. 1996. Successful amplification of PCR-amplified DNA markers for paternity determination in rhesus monkeys (*Macaca mulatta*) and chimpanzees (*Pan troglodytes*). *Lab. Primate Newsl.*, 35, 1–4.

Field, D., Chemnick, L., Robbins, M., Garner, K., & Ryder, O. 1998. Paternity determination in captive lowland gorillas and orangutans and wild mountain gorillas by microsatellite analysis. *Primates*, 39, 199–209.

Fietz, J., Zischler, H., Schwiegk, C., Tomiuk, J., Dausmann, K. H., & Ganzhorn, J. U. 2000. High rates of extra-pair young in the pair-living fat-tailed dwarf lemur, *Cheirogaleus medius. Behav. Ecol. Sociobiol.*, 49, 8–17.

Gagneux, P., Boesch, C., & Woodruff, D. S. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Mol. Ecol.*, 6, 861–868.

Gagneux, P., Boesch, C., & Woodruff, D. S. 1999. Female reproductive strategies, paternity and community structure in wild West African chimpanzees. *Anim. Behav.*, 57, 19–32.

Gagneux, P., Woodruff, D. S., & Boesch, C. 1997. Furtive mating in female chimpanzees. *Nature*, 387, 358–359.

Gerloff, U., Hartung, B., Fruth, B., Hohmann, G., & Tautz, D. 1999. Intracommunity relationships, dispersal pattern and paternity success in a wild living community of Bonobos (*Pan paniscus*) determined from DNA analysis of faecal samples. *Proc. Royal Soc. Lond., B*, 266, 1189–1195.

Gerloff, U., Schlötterer, C., Rassmann, K., Rambold, I., Hohmann, G., Fruth, B., & Tautz, D. 1995. Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living Bonobos (*Pan paniscus*). *Mol. Ecol.*, 4, 515–518.

Goldberg, T. & Wrangham, R. 1997. Genetic correlates of social behaviour in wild chimpanzees: evidence from mitochondrial DNA. *Anim. Behav.*, 54, 559–570.

Goodnight, K. F. & Queller, D. C. 1999. Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Mol. Ecol.*, 8, 1231–1234.

Goossens, B., Latour, S., Vidal, C., Jamart, A., Ancrenaz, M., & Bruford, M. W. 2000. Twenty new microsatellite loci for use with hair and faecal samples in the chimpanzee (*Pan troglodytes troglodytes*). *Folia Primatol.*, 71, 177–180.

Goossens, B., Waits, L. P., & Taberlet, P. 1998. Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Mol. Ecol.*, 7, 1237–1241.

Goossens, B. G., Chikhi, L., Utami, S. S., de Ruiter, J. R., & Bruford, M. W. 2000. Multiple-samples and multiple-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Cons. Genet.*, 1, 157–162.

Greenwood, A. & Pääbo, S. 1999. Nuclear insertion sequences of mitochondrial DNA predominate in hair but not in blood of elephants. *Mol. Ecol.*, 8, 133–137.

Griffiths, R. & Tiwari, B. 1993. Primers for the differential amplification of the sex-determining region Y gene in a range of mammal species. *Mol. Ecol.*, 2, 405–406.

Gut, I. G. 2001. Automation in genotyping of single nucleotide polymorphisms. *Hum. Mutat.*, 17, 475–492.

Hacia, J. G., Fan, J.-B., Ryder, O. A., Jin, L., Edgemon, K., Ghandour, G., Mayer, R. A., Sun, B., Hsie, L., Robbins, C. M., Brody, L. C., Wang, D., Lander, E. S., Lipshutz, R., Fodor, S. P. A., & Collins, F. S. 1999. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat. Genet.*, 22, 164–167.

Hamilton, D. W. 1964. The genetical evolution of social behaviour 1 & 2. *J. Theor. Biol.*, 7, 1–52.

Hashimoto, C., Furuichi, T., & Takenaka, O. 1996. Matrilineal kin relationship and social behavior of wild bonobos: sequencing the D-loop region of mitochondrial DNA. *Primates*, 37, 305–318.

Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F., & Erlich, H. A. 1988. DNA typing from single hairs. *Nature*, 332, 543–546.

Hirschhorn, J. N., Sklar, P., Lindblad-Toh, K., Lim, Y.-M., Ruiz-Gutierrez, M., Bolk, S., Langhorst, B., Schaffner, S., Winchester, E., & Lander, E. S. 2000. SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping. *Proc. Natl. Acad. Sci USA*, 97, 12164–12169.

Hoelzel, A. R. 1998. *Molecular Genetic Analysis of Populations: A Practical Approach.* Oxford: Oxford University Press.

Hoelzel, A. R. & Green, A. 1998. PCR protocols and population analysis by direct DNA sequencing and PCR-based DNA fingerprinting. In: *Molecular Genetic Analysis of Populations* (Ed. by A. R. Hoelzel), pp. 201–236. Oxford: Oxford University Press.

Immel, U. D., Hummel, S., & Herrmann, B. 2000. Reconstruction of kinship by fecal DNA analysis of orangutans. *Anthropol. Anz.*, 58, 63–67.

Inoue, M., Mitsunaga, F., Ohsawa, H., Takenaka, A., Sugiyama, Y., Soumah, A. G., & Takenaka, O. 1992. Paternity testing in captive Japanese macaques (*Macaca fuscata*) using DNA fingerprinting. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 131–140. Basel: Karger.

Inoue, M., Takenaka, A., Tanaka, S., Kominami, R., & Takenaka, O. 1990. Paternity discrimination in a Japanese macaque troop by DNA fingerprinting. *Primates*, 31, 563–570.

Inoue, M. & Takenaka, O. 1993. Japanese macaque microsatellite PCR primers for paternity testing. *Primates*, 34, 37–45.

Jarne, P. & Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.*, 11, 424–429.

Jeffreys, A. J. 1987. Highly variable minisatellites and DNA fingerprints. *Biochem. Soc. Trans.*, 15, 309–317.

Jeffreys, A. J., Turner, M., & Debenham, P. 1991. The efficiency of multilocus DNA fingerprint probes for individualization and establishment of family relationships determined from extensive casework. *Am. J. Hum. Genet.*, 48, 824–840.

Jeffreys, A. J., Wilson, V., & Thein, S. L. 1985. Hypervariable "minisatellite" regions in human DNA. *Nature*, 314, 67–73.

Jekielek, J. & Strobeck, C. 1999. Characterization of polymorphic brown lemur (*Eulemur fulvus*) microsatellite loci and their amplification in the family Lemuridae. *Mol. Ecol.*, 8, 901–903.

Kayser, M., Rit... cation of h... cross-spec...

Keane, B., Ditt... toque mac... *Mol. Ecol.*...

Kocher, T. D.,... Wilson, A... tion and s... 6200.

Krawczak, M. ... *Electropho...*

Kruglyak, L. 1... *Genet.*, 17...

Kwok, P.-Y. 2... 95–100.

Lanave, C., At... C. 1999. U... sequences...

Lander, E. S. 1...

Launhardt, K.,... lites adap... *entellus*)...

Li, C. C., We... chance ar...

Litt, M. & Lut... of a dinu... 397–401.

Luikart, G. &... *Ecol. Ev...*

Lynch, M. 19... 584–599.

Lynch, M. &... *Genetics.*...

Lyons, L. A.,... S. J. 199... mammal...

Marshall, E. ...

Marshall, T. ... likelihoo...

Martin, R. D.... *and The...*

Mhlanga, M. ... polymor...

Mitani, J., M... in wild ...

Moore, S. S.... 1991. T... allows t... 10, 654...

Kayser, M., Ritter, H., Bercovitch, F., Mrug, M., Roewer, L., & Nurnberg, P. 1996. Identification of highly polymorphic microsatellites in the rhesus macaque *Macaca mulatta* by cross-species amplification. *Mol. Ecol.*, 5, 157–159.

Keane, B., Dittus, W. P., & Melnick, D. J. 1997. Paternity assessment in wild groups of toque macaques *Macaca sinica* at Polonnaruwa, Sri Lanka using molecular markers. *Mol. Ecol.*, 6, 267–282.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA*, 86, 6196–6200.

Krawczak, M. 1999. Informativity assessment for biallelic single nucleotide polymorphisms. *Electrophoresis*, 20, 1676–1681.

Kruglyak, L. 1997. The use of a genetic map of biallelic markers in linkage studies. *Nat. Genet.*, 17, 21–24.

Kwok, P.-Y. 2000. High-throughput genotyping assay approaches. *Pharmacogenomics*, 1, 95–100.

Lanave, C., Attimonelli, M., De Robertis, M., Licciulli, F., Liuni, S., Sbisa, E., & Saccone, C. 1999. Update of AMmtDB: a database of multi-aligned metazoa mitochondrial DNA sequences. *Nucleic Acids Res.*, 27, 134–137.

Lander, E. S. 1989. DNA fingerprinting on trial. *Nature*, 339, 501–505.

Launhardt, K., Epplen, C., Epplen, J. T., & Winkler, P. 1998. Amplification of microsatellites adapted from human systems in faecal DNA of wild hanuman langurs (*Presbytis entellus*). *Electrophoresis*, 19, 1356–1361.

Li, C. C., Weeks, D. E., & Chakravarti, A. 1993. Similarity of DNA fingerprints due to chance and relatedness. *Hum. Hered.*, 43, 45–52.

Litt, M. & Luty, J. A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.*, 44, 397–401.

Luikart, G. & England, P. R. 1999. Statistical analysis of microsatellite DNA data. *Trends Ecol. Evol.*, 14, 253–256.

Lynch, M. 1988. Estimation of relatedness by DNA fingerprinting. *Mol. Biol. Evol.*, 5, 584–599.

Lynch, M. & Ritland, K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics*, 152, 1753–1766.

Lyons, L. A., Laughlin, T. F., Copeland, N. G., Jenkins, N. A., Womack, J. E., & O'Brien, S. J. 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nat. Genet.*, 15, 47–56.

Marshall, E. 1997. The hunting of the SNP. *Science*, 278, 2047.

Marshall, T. C., Slate, J., Kruuk, L. E. B., & Pemberton, J. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.*, 7, 639–655.

Martin, R. D., Dixson, A. F., & Wickings, E. J. 1992. *Paternity in Primates: Genetic Tests and Theories*. Basel: Karger.

Mhlanga, M. M. & Malmberg, L. 2001. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods*, 25, 463–471.

Mitani, J., Merriwether, D. A., & Zhang, C. 2000. Male affiliation, cooperation, and kinship in wild chimpanzees. *Anim. Behav.*, 59, 885–893.

Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M., & Hetzel, D. J. S. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, 10, 654–660.

Morin, P. A., Chambers, K. E., Boesch, C., & Vigilant, L. 2001. Quantitative PCR analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes*). *Mol. Ecol.*, 10, 1835–1844.

Morin, P. A., Kanthaswamy, S., & Smith, D. G. 1997. Simple sequence repeat (SSR) polymorphisms for colony management and population genetics in Rhesus macaques (*Macaca mulatta*). *Am. J. Primatol.*, 42, 199–213.

Morin, P. A., Mahboubi, P., Wedel, S., & Rogers, J. 1998. Rapid screening and comparison of human microsatellite markers in baboons: allele size is conserved, but allele number is not. *Genomics*, 53, 12–20.

Morin, P. A., Moore, J. J., Chakraborty, R., Jin, L., Goodall, J., & Woodruff, D. S. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science*, 265, 1193–1201.

Morin, P. A., Moore, J. J., Wallis, J., & Woodruff, D. S. 1994. Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Mol. Ecol.*, 3, 469–478.

Morin, P. A., Moore, J. J., & Woodruff, D. S. 1992. Identification of chimpanzee subspecies with DNA from hair and allele specific probes. *Proc. Royal Soc. Lond., B*, 249, 293–297.

Morin, P. A. & Ryder, O. A. 1991. Founder contribution and pedigree inference in a captive breeding colony of lion-tailed macaques, using mitochondrial DNA and DNA fingerprint analyses. *Zoo Biol.*, 10, 341–352.

Morin, P. A., Saiz, R., & Monjazeb, A. 1999. High-throughput single nucleotide polymorphism genotyping by fluorescent 5′ exonuclease assay. *BioTechniques*, 27, 538–540, 542, 544.

Morin, P. A., Wallis, J., Moore, J. J., Chakraborty, R., & Woodruff, D. S. 1993. Noninvasive sampling and DNA amplification for paternity exclusion, community structure, and phylogeography in wild chimpanzees. *Primates*, 34, 347–356.

Morin, P. A. & Woodruff, D. S. 1992. Paternity exclusion using multiple hypervariable microsatellite loci amplified from nuclear DNA of hair cells. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 63–81. Basel: Karger.

Morin, P. A. & Woodruff, D. S. 1996. Noninvasive sampling for vertebrate conservation. In: *Molecular Approaches in Conservation* (Ed. by R. Wayne & T. Smith), pp. 298–313. Oxford: Oxford University Press.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., & O'Brien, S. J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature*, 409, 614–618.

Navidi, W., Arnheim, N., & Waterman, M. S. 1992. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *Am. J. Hum. Genet.*, 50, 347–359.

Oka, T. & Takenaka, O. 2001. Wild gibbon's parentage tested by noninvasive DNA sampling and PCR-amplified polymorphic microsatellites. *Primates*, 42, 67–73.

Palumbi, S. R. & Baker, C. S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Mol. Biol. Evol.*, 11, 426–435.

Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., & Syvänen, A.-C. 1997. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.*, 7, 606–614.

Pemberton, J. M., Coltman, D. W., Smith, J. A., & Pilkington, J. G. 1999. Molecular analysis of a promiscuous, fluctuating mating system. *Biol. J. Linnean Soc.*, 68, 289–301.

Pesole, G., Gissi, C., De Chirico, A., & Saccone, C. 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *J. Mol. Evol.*, 48, 427–434.

analysis
ld chim-

R) poly-
ues (Ma-

mparison
e number

S. 1994.
Science,

sion in a
ats. Mol.

ubspecies
293–297.
a captive
IA finger-

polymor-
538–540,

oninvasive
cture, and

ervariable
Primates:
kings), pp.

nservation.
, pp. 298–

3rien, S. J.
ature, 409,

or accurate
ations. Am.

DNA sam-

clear intron
5.
sequencing:
ys. Genome

cular analy-
289–301.
tion rate of

Pope, T. R. 1990. The reproductive consequences of male cooperation in the red howler monkey: paternity exclusion in multi-male troops using genetic markers. *Behav. Ecol. Sociobiol.*, 27, 439–446.

Pope, T. R. 2000. Reproductive success increases with degree of kinship in cooperative coalitions of female red howler monkeys (*Alouatta seniculus*). *Behav. Ecol. Sociobiol.*, 48, 253–267.

Queller, D. C. & Goodnight, K. F. 1989. Estimating relatedness using genetic markers. *Evolution*, 43, 258–275.

Queller, D. C., Strassmann, J. E., & Hughes, C. R. 1993. Microsatellites and kinship. *Trends Ecol. Evol.*, 8, 285–288.

Ritland, K. 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. *Genet. Res.*, 67, 175–185.

Rogers, J., Mahaney, M. C., Witte, S. M., Nair, S., Newman, D., Wedel, S., Rodriguez, L. A., Rice, K. S., Slifer, S. H., Perelygin, A., Slifer, M., Palladino-Negro, P., Newman, T., Chambers, K., Joslyn, G., Parry, P., & Morin, P. A. 2000. A genetic linkage map of the baboon (*Papio hamadryas*) genome based on human microsatellite polymorphisms. *Genomics*, 67, 237–247.

Rogers, J., Witte, S. M., & Slifer, M. A. 1995. Five new microsatellite DNA polymorphisms in squirrel monkey (*Saimiri boliviensis*). *Am. J. Primatol.*, 36, 151.

Rosenbaum, H. C. & Deinard, A. S. 1998. Caution before claim: an overview of microsatellite analysis in ecology and evolutionary biology. In: *Molecular Approaches to Ecology and Evolution* (Ed. by R. DeSalle & B. Schierwater), pp. 87–106. Boston: Birkhäuser.

Ross, K. G. 2001. Molecular ecology of social behaviour: analyses of breeding systems and genetic structure. *Mol. Ecol.*, 10, 265–284.

Schlötterer, C. 1998. Microsatellites. In: *Molecular Genetic Analysis of Populations* (Ed. by A. R. Hoelzel), pp. 237–262. New York: Oxford University Press.

Schlötterer, C., Amos, B., & Tautz, D. 1991. Conservation of polymorphic simple sequences in cetacean species. *Nature*, 354, 63–65.

Schlötterer, C. & Pemberton, J. 1998. The use of microsatellites for genetic analysis of natural populations—a critical review. In: *Molecular Approaches to Ecology and Evolution* (Ed. by R. DeSalle & B. Schierwater), pp. 71–86. Berlin: Birkhäuser Verlag.

Schmitz, J., Ohme, M., & Zischler, H. 2000. The complete mitochondrial genome of *Tupaia belangeri* and the phylogenetic affiliation of Scandentia to other eutherian orders. *Mol. Biol. Evol.*, 17, 1334–1343.

Shi, M. M. 2001. Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.*, 47, 164–172.

Shubitowski, D., Venta, P., Douglass, C., Zhou, R.-X., & Ewart, S. 2001. Polymorphism identification within 50 equine gene-specific sequence tagged sites. *Anim. Genet.*, 32, 78–88.

Silk, J. B. 1987. Social behavior in evolutionary perspective. In: *Primate Societies* (Ed. by B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham, & T. T. Struhsaker), pp. 318–329. Chicago: University of Chicago Press.

Smith, D.-G., Kanthaswamy, S., Disbrow, M., & Wagner, J.-L. 1999. Reconstruction of parentage in a band of captive hamadryas baboons. *Int. J. Primatol.*, 20, 415–429.

Smith, D. G., Rolfs, B., & Lorenz, J. 1992. A comparison of the success of electrophoretic methods and DNA fingerprinting for paternity testing in captive groups of rhesus macaques. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 32–52. Basel: Karger.

Smith, K. L., Alberts, S. C., Bayes, M. K., Bruford, M. W., Altmann, J., & Ober, C. 2000. Cross-species amplification, noninvasive genotyping, and non-Mendelian inheritance of human STRPs in Savannah baboons. *Am. J. Primatol.*, 51, 219–227.

Smith, T. B. & Wayne, R. 1996. *Molecular Genetic Approaches in Conservation.* New York: Oxford University Press.

Stone, A. C., Griffiths, R. C., Zegura, S. L., & Hammer, M. F. 2002. High levels of Y-chromosome nucleotide diversity in the genus Pan. *Proc. Natl. Acad. Sci. USA*, 99, 43–48.

Sullivan, K. M., Mannucci, A., Kimpton, C. P., & Gill, P. 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *BioTechniques*, 15, 637–641.

Sunnucks, P. 2000. Efficient genetic markers for population biology. *Trends Ecol. Evol.*, 15, 199–203.

Syvanen, A. C. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.*, 2, 930–942.

Syvänen, A.-C. & Landegren, U. 1994. Detection of point mutations by solid-phase methods. *Hum. Mutat.*, 3, 172–179.

Taberlet, P., Camarra, J.-J., Griffin, S., Uhrés, E., Hanotte, O., Waits, L. P., Dubois-Paganon, C., Burke, T., & Bouvet, J. 1997. Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Mol. Ecol.*, 6, 869–876.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P., & Bouvet, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res.*, 24, 3189–3194.

Taberlet, P., Waits, L. P., & Luikart, G. 1999. Noninvasive genetic sampling: look before you leap. *Trends Ecol. Evol.*, 14, 323–327.

Takasaki, H. & Takenaka, O. 1991. Paternity testing in chimpanzees with DNA amplification from hairs and buccal cells in wadges: a preliminary note. In: *Primatology Today: Proceedings of the XIIIth Congress of the IPS, Nagoya and Kyoto* (Ed. by M. Iwamoto), pp. 613–616. Amsterdam: Elsevier.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.*, 17, 6463–6471.

Thomas, M. G., Bradman, N., & Flinn, H. M. 1999. High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. *Hum. Genet.*, 105, 577–581.

Thomas, M. G., Parfitt, T., Weiss, D. A., Skorecki, K., Wilson, J. F., Le Roux, M., Bradman, N., & Goldstein, D. B. 2000. Y-chromosome traveling south: the Cohen modal haplotype and the origins of the Lemba—the Black Jews of Southern Africa. *Am. J. Hum. Genet.*, 66, 674–686.

Trivers, R. L. 1985. *Social Evolution.* Menlo Park: Benjamin/Cummings.

Turner, T. R., Weiss, M. L., & Pereira, M. E. 1992. DNA fingerprinting and paternity assessment in Old World monkeys and ringtailed lemurs. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 96–112. Basel: Karger.

Underhill, P. A., Jin, L., Lin, A. A., Mehdi, S. Q., Jenkins, T., Vollrath, D., Davis, R. W., Cavalli-Sforza, L. L., & Oefner, P. J. 1997. Detection of numerous Y-chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.*, 7, 996–1005.

Van De Casteele, T., Galbusera, P., & Matthysen, E. 2001. A comparison of microsatellite-based pairwise relatedness estimators. *Mol. Ecol.*, 10, 1539–1549.

Venta, P., Brouillette, J., Yuzbasiyan-Gurkan, V., & Brewer, G. 1996. Gene-specific universal mammalian sequence tagged sites: application to the canine genome. *Biochem. Genet.*, 34, 321–341.

Vigilant, I from

Vigilant, I chim

von Seges free-

Warren, K Heen *Pong

Weatherh finge

Weber, J. *mics*

Weber, J. be ty

Webster, birds (Ed.

Wickings of C R. D

Wilson, J 379,

Wise, C. and 707

Wrangha 75,

Zane, L. revi

Zhang, ing prir

Zischler ass

Zischler the que

Vigilant, L. 1999. An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. *Biol. Chem.*, 380, 1329–1331.

Vigilant, L., Hofreiter, M., Siedel, H., & Boesch, C. 2001. Paternity and relatedness in wild chimpanzee communities. *Proc. Natl. Acad. Sci. USA*, 98, 12890–12895.

von Segesser, F., Scheffrahn, W., & Martin, R. D. 1995. Parentage analysis within a semi-free-ranging group of Barbary macaques *Macaca sylvanus. Mol. Ecol.*, 4, 115–120.

Warren, K. S., Verschoor, E. J., Langenhuijzen, S., Heriyanto, Swan, R. A., Vigilant, L., & Heeney, J. L. 2001. Speciation and intrasubspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus. Mol. Biol. Evol.*, 18, 472–480.

Weatherhead, P. J. & Montgomerie, R. D. 1991. Good news and bad news about DNA fingerprinting. *Trends Ecol. Evol.*, 6, 173–174.

Weber, J. L. 1990. Informativeness of human (dC-dA)n.(dT-dG)n polymorphisms. *Genomics*, 7, 524–530.

Weber, J. L. & May, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.*, 44, 388–396.

Webster, M. S. & Westneat, D. F. 1998. The use of molecular markers to study kinship in birds: techniques and questions. In: *Molecular Approaches to Ecology and Evolution* (Ed. by R. DeSalle & B. Schierwater), pp. 7–35. Berlin: Birkhäuser Verlag.

Wickings, E. J. & Dixson, A. F. 1992. Application of DNA fingerprinting to familial studies of Gobonese primates. In: *Paternity in Primates: Genetic Tests and Theories* (Ed. by R. D. Martin, A. F. Dixson, & E. J. Wickings), pp. 113–130. Basel: Karger.

Wilson, J. F. & Erlandsson, R. 1998. Sexing of human and other primate DNA. *Biol. Chem.*, 379, 1287–1288.

Wise, C. A., Michaela, S., Rubinsztein, D. C., & Easteal, S. 1997. Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees. *Mol. Biol. Evol.*, 14, 707–716.

Wrangham, R. W. 1980. An ecological model of female-bonded primate groups. *Behaviour*, 75, 262–300.

Zane, L., Bargelloni, L., & Patarnello, T. 2001. Strategies for microsatellite isolation: A review. *Mol. Ecol.*, 11, 1–16.

Zhang, Y. W., Morin, P. A., Ryder, O. A., & Zhang, Y. P. 2001. Microsatellite loci screening in gorillas (*Gorilla gorilla gorilla*) and orangutans (*Pongo pygmaeus*) using human primers prescreened on baboons. *Cons. Genet.*, 2, 391–395.

Zischler, H. 2000. Nuclear integrations of mitochondrial DNA in primates: inference of associated mutational events. *Electrophoresis*, 21, 531–536.

Zischler, H., Geisert, H., & Castresana, J. 1998. A hominoid-specific nuclear insertion of the mitochondrial D-loop: implications for reconstructing ancestral mitochondrial sequences. *Mol. Biol. Evol.*, 15, 463–469.