ANIMAL BEHAVIOUR, 2008, **75**, 1173–1180
doi:10.1016/j.anbehav.2007.08.027

COMMENTARIES

Can't get there from here: inferring kinship from pairwise genetic relatedness

RUSSELL C. VAN HORN*, **JEANNE ALTMANN**†‡ & **SUSAN C. ALBERTS***‡
*Department of Biology, Duke University
†Department of Ecology and Evolutionary Biology, Princeton University
‡Institute of Primate Research, National Museums of Kenya

*(Received 17 May 2007; initial acceptance 7 June 2007;
final acceptance 26 August 2007; published online 28 January 2008; MS. number: AS-1316R)*

Keywords: genetic similarity; kinship; *Papio cynocephalus*; pedigree; relatedness; savannah baboon

Kinship among animals has important implications for their behaviour, as it determines their opportunities for nepotism and their risk of inbreeding (e.g. Hamilton 1963; Wade 1979; Trontti et al. 2005). The most accurate way to assess kinship is through complete knowledge of pedigrees (we use the term kinship to refer to categorical levels of coancestry, as would be derived from pedigree construction). However, precisely identifying kinship among all individuals in a population is rarely, if ever, possible from observations alone, and researchers often either use genetic data alone (see reviews in Blouin 2003; Garant & Kruuk 2005) or use genetic data to complement their behavioural data (e.g. Taylor et al. 1997; Kappeler et al. 2002; Kruuk et al. 2002; Carpenter et al. 2005; Cutrera et al. 2005; Randall et al. 2007).

Several methods are available by which genetic data can be used to estimate relatedness, as a continuous estimate of allele sharing due to coancestry. However, each of these methods (e.g. Queller & Goodnight 1989; Lynch & Ritland 1999; Wang 2002) has inherent limitations (see discussions in van de Casteele et al. 2001; Thomas et al. 2002; Wang 2002; Blouin 2003; Thomas 2005; Fernández & Toro 2006). In particular, the accuracy and precision of these genetic estimators of relatedness are known to be influenced by characteristics of the markers and the sample of individuals under investigation: the number of loci, the number of alleles per locus, the distribution of alleles within loci, and the distribution of coancestry within the sample and the population (see reviews in Lynch & Ritland 1999; van de Casteele et al. 2001; Toro et al. 2002; Blouin 2003). These limitations may lead to erroneous conclusions about kinship if researchers rely too heavily on genetic similarity (see discussion in Blouin et al. 1996; Toro et al. 2002; Russello & Amato 2004), rather than focusing on pedigree construction (Thomas et al. 2002; Toro et al. 2002). As most recently shown by Csilléry et al. (2006), dyads of known pedigree kinship are often misclassified solely on the basis of genetic estimates of relatedness, even when using the estimator that should have performed best, based on simulated analyses. Csilléry et al. (2006) illustrated that five natural populations with various mating systems consistently included far fewer closely related dyads than assumed in prior simulations of estimator performance, so those assessments of estimator performance were overly optimistic. The consistently high misclassification rates observed by Csilléry et al. (2006), and the overall tendency for dyads to be classified as closer kin than they really were, may leave the reader feeling somewhat less than optimistic about the use of genetic markers to infer kinship between pairs of individuals.

However, genetic markers are often used to infer kinship because there do not seem to be any alternatives. For example, full pedigrees are not known for all giant pandas, *Ailuropoda melanoleuca*, in captivity, yet those directing captive-breeding efforts must somehow prioritize mating between pairs of pandas (David et al. 2006). We suspect that researchers often implicitly operate under the principle that 'some data are better than no data'. Whether some data are better than no data (i.e. whether some

Correspondence and present address: R. C. Van Horn, CRES-Zoological Society of San Diego, Giant Panda Conservation Unit, P.O. Box 120551, San Diego, CA 92112-0551, U.S.A. (email: rvanhorn@sandiegozoo.org). J. Altmann is at the Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, U.S.A. S. C. Alberts is at the Department of Biology, Duke University, P.O. 90338, Durham, NC 27708, U.S.A.

0003–3472/08/$32.00/0

data tempt researchers to draw erroneous conclusions that they would not attempt to draw in the absence of data) depends on the consequences of an erroneous conclusion, but given that genetic estimators of relatedness are blunt instruments, as clearly indicated by Csilléry et al. (2006) and others, how should researchers proceed?

Our comments apply in varying degrees to any genetic estimator of relatedness, none of which is consistently better than others at identifying pairwise relatedness (see van de Casteele et al. 2001; Toro et al. 2002; Thomas 2005). For illustration, we used Queller & Goodnight's (1989) genetic estimate of pairwise relatedness, *R*, which is the mostly widely used estimator in behavioural research. Queller & Goodnight's *R*, and other established genetic estimators of relatedness, have two primary disadvantages. First, *R* values are variable to a much greater extent than can be accounted for by true variance in the actual pedigree *r* values. *R* between parents and offspring may vary substantially (e.g. Kleven et al. 2005) even though there is no variation in pedigree *r* within parent–offspring pairs (i.e. $r = 0.5$, barring inbreeding), and the variation in *R* among more distant kin is much greater than the possible true variation in allele sharing within classes of distant kin (Guo 1996; Guo & Xiong 1997). To obtain precise estimates of *R* with low variance may require tens or hundreds of autosomal microsatellite loci (Toro et al. 2002; Csilléry et al. 2006; Kalinowski et al. 2006). Maximum likelihood methods are often used to evaluate the meaningfulness of genetic similarity measures like *R* (e.g. Goodnight & Queller 1999), but even these methods require large numbers of loci to discriminate dyads with $r = 0.5$ from dyads with $r = 0.25$ or $r = 0$ (Blouin 2003). Few, if any, investigators currently working with wild animals are able to generate the volume of data required to precisely assess kinship with the current genetic estimators of relatedness; this is particularly true if DNA is obtained through noninvasive sources.

The second disadvantage of all pairwise estimators of kinship, including *R*, is that they may produce results that are inconsistent with logical pedigree structure (Fernández & Toro 2006). For example, a given pairwise estimator might classify individuals X and Y as full siblings, and Y and Z as half-siblings; in this case, logic demands that X and Z must be half-siblings, but the marker-based estimator may classify X and Z as something other than half-siblings (see discussion in Fernández & Toro 2006). This problem only affects closely related sets of dyads (i.e. among three individuals, where $r = 0.5$ for one dyad and $r \geq 0.25$ for a second dyad) but the resulting inconsistencies may be considerable in these cases. If a researcher with only genetic data inferred kinship from relatedness values and observed a logical inconsistency in kinship assignment (e.g. both dyads X–Y and Y–Z classed as full sibs, but dyad X–Z classed as half-sibs), could she resolve this contradiction?

Pairwise relatedness estimators were not originally developed to classify kinship between pairs of individuals (see Blouin et al. 1996), yet researchers use them for this purpose because they have few alternatives. Although imprecision in assessing kinship will be less of a problem when examining coarse-grained questions (e.g. How does kinship within social groups compare to that among social groups?), imprecision will be a major problem in addressing questions that require knowledge of kinship within specific dyads. These questions are crucial for both basic research (e.g. Does a female form stronger social bonds with those age peers that are her close kin than with her other age peers?) and for applied research (e.g. To minimize the risks of inbreeding in a captive population, should female A be mated with male B, or with male C?).

Our first goal was to assess the relatedness composition of a natural population of savannah baboons, *Papio cynocephalus*. Our second goal was to estimate the rate at which baboon dyads would have their kinship misclassified from relatedness estimates (as done by Csilléry et al. 2006 for several other populations of wild mammals), and from a commonly used likelihood approach (i.e. KINSHIP; Goodnight & Queller 1999). Our third goal was to stress that improvements in pairwise genetic estimators of relatedness and kinship (e.g. Milligan 2003; Kalinowski et al. 2006) do not obviate the need for logical consistency when interpreting pairwise kinship (see Fernández & Toro 2006). Our fourth goal was to suggest how biologists might be able to address some of their research questions by using assigned ranges of kinship derived from a combination of pedigree and demographic data.

Methods

We present data from savannah baboons in and around Amboseli National Park, Kenya (see details in Van Horn et al. 2007). For all females and many males in this population, we had extensive knowledge of maternal kinship, in some cases extending back for six (but more often two to five) generations, based on observed pregnancies, births, and mother–offspring relationships. Maternity is essentially detected without error; pregnancy is detected months before parturition based on female physiological signals and cues, and due dates are estimated from these data (see Beehner et al. 2006 for details). In addition, asynchrony of births and the census of each social group at least once every 2.5 days make it extremely unlikely that maternal kin are misidentified; our genetic parentage analyses have confirmed our lack of error in identifying maternity. However, these data did not allow us to identify paternal kin or distant maternal kin; for many pairs of animals we had only partial pedigrees. For instance, we might know that two females were maternal first cousins (i.e. their mothers were maternal sisters), but we might not know their paternal relatedness.

Genotyping and paternity assignment

Maternal kin were identified after we determined maternity from long-term field observations (e.g. Altmann & Alberts 2005; and see above) for 809 baboons. In contrast, paternal kin were identified after we assigned paternity for 340 baboons using multilocus genotypes and information about which mature males were present in the population when offspring were conceived. To assign paternity we genotyped 504 Amboseli baboons at multiple microsatellite loci ($\overline{X} \pm SD = 13.81 \pm 0.78$ loci per individual), genotyping most animals at 14 loci (459 of 504, 91.1%) as described in Alberts et al. (2006). All loci were in Hardy–Weinberg equilibrium, with a mean ± SD of 11.07 ± 2.616 alleles per locus. Mean observed heterozygosity across loci was 0.813 ± 0.039. Males were considered potential fathers if they were present in the population and had enlarged testicles, a presumed indication of sperm production (Castracane et al. 1986; Alberts & Altmann 1995), when the offspring was conceived. Paternal genotypes were inferred from the filial and maternal genotypes (known for most offspring). A male was assigned as father if he was the only male with no genetic mismatches with a given offspring, or he had a single genetic mismatch with the offspring and the paternity assignment was supported at the 95% confidence level by the maximum likelihood method employed in CERVUS 2.0 (Marshall et al. 1998). To generate confidence levels with CERVUS, 10 000 simulations were run with the following parameters: 70% of candidate parents (62) sampled, 98% of loci typed, and a typing error rate of 0.01. The typing error rate was set as the rate of mismatches between mothers and offspring. As our noninvasive sampling efforts have improved over time, we have been able to assign paternity to virtually all sampled offspring (see Alberts et al. 2006 for details). Offspring for which we have not assigned paternity are those that have died (usually due to predation) before we obtained samples; we assume that these are a random sample with respect to microsatellite genotypes.

Assessing the degree to which relatedness corresponds to kinship

We used known maternal kin relationships and assigned paternities to determine total pedigree coefficients of relatedness, *r*, for dyads of baboons. These *r* values would accurately reflect true kinship if full pedigrees were known for all animals. However, from dyads for which we had incomplete pedigrees (e.g. when fathers or grandfathers were not known for one or both animals; Fig. 1), our *r* values were the minimum possible pedigree *r* values within a dyad: a dyad may have been more, but not less, closely related than indicated by our pedigree *r* value (Dittus 1988; Jones et al. 2002).

We next used RELATEDNESS 5.0 to estimate relatedness with Queller & Goodnight's (1989) genetic estimate of pairwise relatedness, *R*. We then assessed the precision and accuracy of *R* as an estimator of pedigree kinship in Amboseli baboons by creating reference distributions of *R* from dyads for which all four grandparents had been identified ($N = 2913$ dyads of males and females of all ages, excluding 13 dyads of complex kinship; e.g. if they were both half-sibling and half-cousin; Fig. 2). Unless we knew otherwise we assumed that the grandparents in the pedigree were unrelated. If this assumption was incorrect and two grandparents were related at $r = 0.5$ (i.e. the highest probable level), then we would underestimate the level of kinship in the focal dyad (e.g. the grandoffspring) by $r = 0.03125$. This imprecision should create only minor noise in the reference distributions. Because correlational analyses are sometimes used to assess whether there is a behavioural response to varying levels of relatedness, we next assessed the strength of the relationship between estimated *R* and pedigree *r* in the reference dyads with Mantel's test as implemented in the software PERMUTE! 3.4, which uses permutation methods as described in Legendre et al. (1994); we ran 999 permutations of the matrices.

To illustrate how *R* values might have led to erroneous conclusions regarding kinship among dyads, we calculated the rates at which we would have misclassified dyads of different kinship types (i.e. unrelated, half-sibling, full

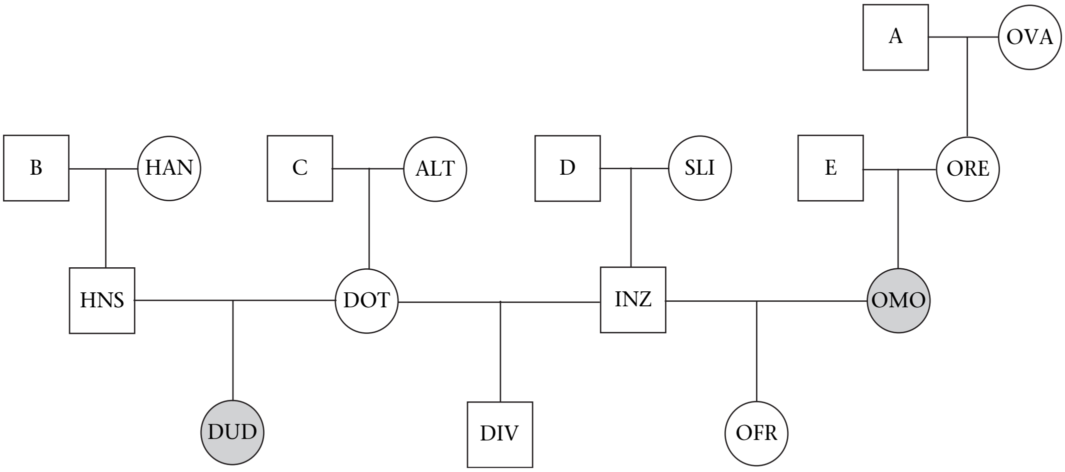

Figure 1. Examples of partial baboon pedigrees based on observed maternity and assigned paternity, illustrating the use of genetic and demographic data to define the possible range of kinship for a dyad. Circles indicate females, squares indicate males and letters indicate unknown individuals. Individuals for whom parents are not indicated are of unknown parentage. In the example, INZ fathered both DIV (son of DOT) and OFR (daughter of OMO). It is known from parentage analysis and demographic data that C ≠ E, and that HNS and INZ are not B, C, D, or E. To assess the kinship between DUD and OMO (shaded), we note that they are unrelated unless B or C = E, in which case OMO is DUD's paternal aunt. B or C and E are the closest possible shared kin between DUD and OMO, and thus the possible range of kinship for DUD and OMO is $0 \le r \le 0.125$. See text for details.

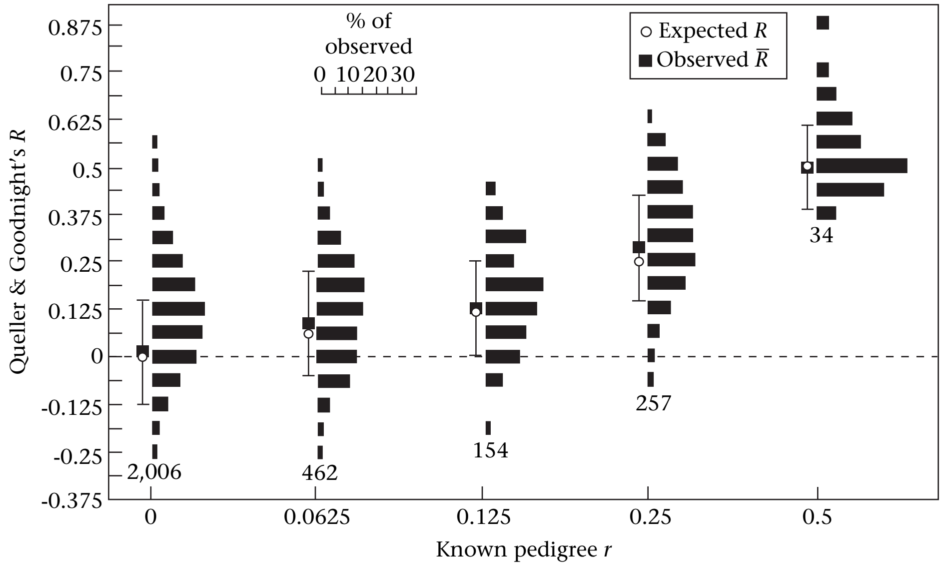

Figure 2. Queller & Goodnight's R values for 2913 dyads of 77 baboons with known parents and grandparents. Open circles represent expected R for each kinship class (e.g. $R = 0$ for nonkin, $R = 0.0625$ for maternal cousins, etc.); filled boxes with error bars represent observed $\bar{R} \pm$ SD, alongside the distribution of R. Sample sizes are the number of dyads for each pedigree class. Thirteen additional dyads are not shown, because they fell into rare classes of kinship (e.g. $r = 0.3125$).

sibling, and parent–offspring), based solely on R values. We did so using two classification procedures. We first generated the distribution of R values from reference dyads of each kin type (e.g. half-sibling), excluding any dyads that represented multiple kin types via complex pedigree (e.g. half-sibling and half-cousin), and we used the midpoints between the means of these distributions as the cutoff points in kinship classification (Blouin et al. 1996). For example, there were 251 half-sibling dyads and six dyads of full aunt or uncle with niece or nephew ($r = 0.25$ for both types of dyads) in the reference data set, with an average R value ($\bar{R}$) of 0.285, and there were 2006 unrelated dyads ($r = 0$) in the reference data set, with $\bar{R} = 0.015$ (Fig. 2). The midpoint between the means of these two distributions is $R = 0.15$, so a dyad with $R = 0.115$ would be classified as an unrelated dyad, rather than as a half-sibling dyad.

We also classified all dyads into kin types with the likelihood method of Goodnight & Queller (1999), using KINSHIP 1.2. To do so we iteratively tested as null hypotheses incrementally closer degrees of kinship between baboons. That is, we first determined whether we could reject the null hypothesis that dyads were unrelated, versus the alternative hypothesis that they were, for example, half-siblings; we then determined whether we could reject the null hypothesis that dyads were half-siblings versus the alternative hypothesis that they were full siblings, and so on. In a typical investigation of kinship between individuals, researchers use pairwise genetic relatedness estimates to assess whether they can reject lower levels of kinship (e.g. unrelated individuals, pedigree $r = 0$) in favour of higher levels of kinship (e.g. half-siblings, pedigree $r = 0.25$). Similarly, to each dyad we assigned the closest degree of kinship not rejected, evaluating statistical significance at $P = 0.05$ by comparing the observed likelihood ratios to those from 10 000 simulated pairs of individuals. After assigning kinship to all dyads via both the R-midpoint method (Blouin et al. 1996) and via KINSHIP (Goodnight & Queller 1999), we then calculated the misclassification rates as the proportion of reference dyads of each kin type that would have been misclassified as another kin type.

Assessing the occurrence of logical inconsistencies in assigned pairwise kinship

All estimators of pairwise kinship, including those based on Queller & Goodnight's (1989) R, may produce results that are inconsistent with logical pedigree structure (Fernández & Toro 2006). For example, in Fig. 1 we know from our pedigree and demographic data that DOT is the mother of DUD and DIV by two different males (HNS and INZ, respectively), and that OMO is the mother of OFR by INZ. DOT has $r = 0.5$ with each of her offspring DUD and DIV, and the dyad DIV–DUD has $r = 0.25$. However, if we had no knowledge of the actual pedigrees of these animals and assessed the kinship of dyads DOT–DUD, DIV–DOT and DIV–DUD independently using KINSHIP, then DOT–DUD (Queller & Goodnight's $R = 0.503$) and DIV–DOT ($R = 0.497$) would be classed as parent–offspring dyads, whereas DIV–DUD ($R = 0.05$) would be illogically classed as a pair of unrelated baboons. The difference between DIV–DUD's classification from KINSHIP (unrelated) and the relationship that they must logically have if DOT has a parent–offspring relationship with both DIV and DUD (DIV–DUD must be full sibs, half-sibs, or grandparent–offspring, with $r = 0.5$ or $r = 0.25$, again ignoring our known pedigree relationships for these baboons) is at least $r = 0.25$.

To quantify the inconsistencies of this sort that would result from using marker-based estimators alone (with no knowledge of actual pedigrees), we used KINSHIP to assign kin relationships to the 2913 reference dyads, and then drew a random sample of 10 reference dyads. Using their assigned kin relationships, we then constructed an inferred pedigree around these 10 reference dyads, and thereby identified 45 other reference dyads that would be linked to them if the KINSHIP assignments were correct (ignoring the known pedigree relationships). For these 45 additional dyads we then compared the kin relationships that would be required by the logic of the inferred pedigree with the kin relationships assigned independently without regard to the inferred pedigree. That is, we compared the fact that DUD and DIV (above) should be full sibs, half-sibs or grandparent–grandoffspring based on their links through DOT, with their assigned pairwise kinship based on KINSHIP (unrelated), and we quantified the difference in units of *r*.

We have focused on the determination of discrete categorical values of kinship. Some research questions may require exact kinship values given a reference population (e.g. Does this particular dyad consist of half-siblings?), but for other questions it may be enough to know whether a particular dyad, or set of dyads, might fall within a specific range of kinship (e.g. Are members of this particular dyad less related than are half-siblings?). We know of no way to use a maximum likelihood approach to incorporate genetic, pedigree and demographic data into an estimate of pedigree kinship, so we suggest instead that other demographic data be used with incomplete pedigree data to define the range of possible kinship for a dyad. The known pedigree kinship between individuals in a dyad establishes their minimum *r* value. For example, between OMO and DUD in Fig. 1, not known to be kin, the minimum pedigree kinship would be $r = 0$: OMO and DUD might truly be unrelated. The maximum possible pedigree kinship within each dyad may be approximated as what their *r* value would be if they shared their nearest unknown kin; DUD and OMO would have $r = 0.125$ if B or C = E (i.e. OMO would be DUD's paternal aunt). The possibility that two individuals share their unknown kin may be evaluated by reference to demographic data. That is, were there any mature males present for the conceptions of both HNS and OMO?

Results

Close kin (i.e. highly related individuals) were uncommon in the 2913 reference dyads (i.e. dyads with known pedigree *r*), regardless of the criterion used to define them. Even if we pooled kin types together by pedigree *r* value, few dyads would be considered close kin; among the reference dyads, 1.2% ($N = 34$) had $r = 0.5$ (i.e. parent–offspring, full sibling), 8.8% ($N = 257$) had $r = 0.25$ (e.g. half-sibling, or full aunt or uncle with niece or nephew), and 68.9% ($N = 2006$) had $r = 0.0$. Mean *R* values ($\overline{R}$) for the reference distributions of dyads with known pedigrees matched the respective pedigree *r* values, but for each known class of kin, our estimates of *R* varied widely (Fig. 2). Consequently, if pedigrees had been incomplete, using *R* alone as a surrogate for pedigree *r* would have produced many errors in kinship assessment (Table 1). Although pedigree *r* within the reference dyads was a statistically significant predictor of *R*, as expected, the relationship was far from perfect ($R^2 = 0.314$, $b = 0.6$, $N = 2913$ dyads, $P = 0.001$). Thus, using *R* as a proxy for *r* in correlational analyses of behaviour would be unlikely to produce valid conclusions.

Based on midpoints of distributions of *R* values, and considering only dyads of four kin classes (i.e. parent–offspring, full sibling, half-sibling and unrelated, $N = 2291$ dyads), we correctly assigned kin type to most of these dyads (79.9% or 1831 dyads using the *R*-midpoint method). However, this method performed best on the unrelated class, which included most dyads (87.5%, or 2006 of 2291 dyads). Therefore, the pooled success rate misrepresents the probability of correctly identifying kin. In fact, of the 285 parent–offspring pairs, full sibling pairs and half-sibling pairs, only 158 dyads (55.4%) were correctly assigned. In particular, dyads were often misassigned to higher categories of kinship. For example, of the 45 dyads classed as parent–offspring pairs by the *R*-midpoint method (Table 1, column 1), 84.4% were not parent–offspring pairs; 10 dyads classed as parent–offspring pairs (22.2%) were in

Table 1. Classification and misclassification of reference dyads ($N = 2291$) of baboons of known pedigree kinship, based on the distribution of their pairwise *R* values (*R*), and by the software KINSHIP (*K*)

	Kinship as classified from genotypic data									
	Parent–offspring		*Full siblings*		*Half-siblings*		*Unrelated*		*Success rate (%)*	
True kinship	*R*	*K*	*R*	*K*	*R*	*K*	*R*	*K*	*R*	*K*
Parent–offspring	**7**	**11**	4	0	0	0	0	0	63.6	100
Full siblings	12	14	**7**	**9**	4	0	0	0	30.4	39.1
Half-siblings	16	13	49	159	**144**	**12**	42	67	57.4	4.8
Unrelated	10	1	0	138	318	17	**1678**	**1850**	83.6	92.2
Success rate (%)	15.6	28.2	11.7	2.9	30.9	41.4	97.6	96.5	**79.9**	**82.1**

These dyads include only baboons for which all parents and grandparents were known. Correctly classified dyads are listed on the diagonal in bold font. For instance, among all dyads that were classified as parent–offspring according to *R* or *K*, only 15.6% or 28.2% (respectively) were actually parent–offspring pairs, as indicated in the bottom row, 'Success rate'. Similarly, among all true parent–offspring pairs, 63.6% and 100% (respectively) were correctly identified as such using *R* or *K* as indicated in the rightmost column, also labelled 'Success rate'.

fact unrelated. Using KINSHIP to assign kinship produced even higher rates of error among close kin. For example, of the 306 pairs that were identified as full siblings by KINSHIP, only nine pairs (2.9%) were actually full siblings (Table 1, column 4). Across the three kin classes (i.e. parent–offspring, full siblings and half-siblings) only 32 of 285 dyads (11.2%) were correctly assigned (Table 1). Although the overall success rate using KINSHIP was moderately high (82.1% of dyads were assigned to the correct kin type; Table 1), this again reflected better performance for truly unrelated pairs, the largest component of our sample, and dyads classified as close kin were often truly unrelated (Table 1).

Assigning kin relationships to dyads based on Queller & Goodnight's *R*, without considering the logic of inferred pedigree structure led to multiple inconsistencies in pedigree structure of the kind described in our earlier example (in which KINSHIP assigned parent–offspring relationships to DOT–DUD and DIV–DOT but classified DIV–DUD as unrelated). Specifically, for 30 of the 45 linked dyads (e.g. cases similar to DIV–DUD), the kin relationships demanded by logic were inconsistent with the kin relationships assigned by KINSHIP, and the magnitude of the mean ± SD difference was $r = 0.317 \pm 0.112$ ($N = 30$). Similar inconsistencies also result from the use of the *R*-midpoint method (data not shown). These logical inconsistencies illustrate an additional disadvantage, beyond the simple misclassification of dyads, that plagues any pairwise assessment of coancestry, particularly when based on a limited number of genetic markers.

The use of demographic data in conjunction with incomplete genetic and pedigree data allows us to assess the possible range of kinship in some dyads. OMO and DUD in Fig. 1 have $R = 0.2965$, and would be assigned $r = 0.5$ via KINSHIP. However, we know from pre-established pedigree ties (i.e. Fig. 1) that $r < 0.5$ for DUD–OMO. From their known pedigree, we do not know the exact kinship between DUD–OMO; they could have $r = 0$ if B or C $\neq$ E (i.e. DUD and OMO are nonkin) or $r = 0.125$ if B or C = E (i.e. if OMO is DUD's paternal aunt). In this case, does demographic data shed any light on the kinship between DUD and OMO? That is, were there any mature males present for the conceptions of both HNS and OMO? From our demographic data we confirmed that there were individual mature males present for the conceptions of both HNS and OMO. Thus, the possible kinship range of DUD–OMO is $0 \leq r \leq 0.125$. For this dyad, demographic data did not narrow the possible range of kinship within the bounds already set by pedigree data, but for many dyads demographic data would further constrain the possible range of kinship set by pedigree data. In addition, if the research question of interest required us to identify dyads that might be full siblings, our demographic data or our pedigree data would lead us to exclude DUD–OMO from our sample even though KINSHIP indicated otherwise.

Discussion

Most baboon reference dyads were not close kin, as was also the situation in all five species examined by Csilléry et al. (2006; meerkat, *Suricata suricata*; great reed warbler, *Acrocephalus arundinaceus*; bighorn sheep, *Ovis canadensis*; red deer, *Cervus elaphus*; Soay sheep, *O. aries*). The rarity of close kin implies that past simulations of the efficiency of relatedness estimators may often not reflect true performance of the estimators (see discussion in Csilléry et al. 2006). Although mean *R* values were as expected from known pedigree relationships, *R* values for individual dyads did not conform well to known pedigree relationships (see also Toro et al. 2002; Csilléry et al. 2006). By using *R* alone, or by using KINSHIP, we could have identified a set of putative unrelated dyads that included very few dyads that were actually closely related. However, any putative set of closely related dyads would have included a large proportion of incorrectly classified dyads. This pattern of kinship misclassification would almost inevitably lead to erroneous conclusions about kinship and behaviour. The asymmetrical bias in misclassification probably reflects both the fact that unrelated dyads are the most common segment of the population and that dyads may be genetically similar because of shared identity by state but not identity by descent. As also seen in other contexts (Thomas et al. 2002; Fernández et al. 2005), molecular data alone were not as informative as pedigree information derived from observations of maternity and genetic assignments of paternity. The maximum likelihood estimator of Milligan (2003), as implemented in Kalinowski et al. (2006), may be a methodological improvement, but it, like all other pairwise estimators of relatedness and kinship, may produce logical inconsistencies. We know of no published method to incorporate the inherent logic of pedigree structure into kinship assessment other than to construct pedigrees. However, the simulated annealing method of Fernández & Toro (2006) holds promise that genetic data may be used to infer logically consistent kinship values. We therefore concur with others (e.g. Garant & Kruuk 2005) that genetic data are often best used to identify pedigree relationships, not replace them. Genetic data from sex-linked loci, in addition to the autosomal loci that are more commonly used, might further refine kinship assessment. Regardless of the loci used, if researchers can avoid depending entirely on genetic data and instead combine genetic data with even limited pedigree and demographic information, our understanding of the role of kinship in behaviour will be enhanced. If an accurate but imprecise assignment of kinship, via genetic relatedness and demographic data, is sufficient to address the research questions of interest, then to us such an approach seems preferable to relying on genotypic data to assign precise but inaccurate levels of kinship. Although categorical measures of relatedness (i.e. pedigree kinship) provide only approximations of the underlying relatedness between two individuals (e.g. Thomas 2005; Csilléry et al. 2006), it seems possible, if not probable, that animals respond to cues of categorical kinship (e.g. nursing associations). The degree to which animal behaviour is influenced by those cues, relative to cues of true genetic relatedness as estimated with current methodology (e.g. Parrott et al. 2007), is yet to be determined.

This work was supported by National Science Foundation (NSF) grant BCS-0323553, and grant NSF IBN-0322613 and its predecessors. For their assistance we thank the Office of the President, Republic of Kenya, the Kenya Wildlife Services, its Amboseli staff and Wardens, the Institute of Primate Research, the National Museums of Kenya, and the members of the Amboseli–Longido pastoralist communities. Particular thanks go to the Amboseli fieldworkers who contributed to sample and data collection, especially R. Mututua, S. Sayialel and J. Warutere. J. Buchan identified many paternal relationships, and helpful comments were provided by E. Archie, R. Davidson, A. Widdig and members of the Alberts laboratory. All protocols were noninvasive and comply with relevant regulations in Kenya (Kenya Research Permit MOEST 13/001/33C 179) and the United States (Duke University IACUC protocol A180-03-05).

References

Alberts, S. C. & Altmann, J. 1995. Preparation and activation: determinants of age at reproductive maturity in male baboons. *Behavioral Ecology and Sociobiology*, **36**, 397–406.

Alberts, S. C., Buchan, J. C. & Altmann, J. 2006. Sexual selection in wild baboons: from mating opportunities to paternity success. *Animal Behaviour*, **72**, 1177–1196.

Altmann, J. & Alberts, S. C. 2005. Growth rates in a wild primate population: ecological influences and maternal effects. *Behavioral Ecology and Sociobiology*, **57**, 490–501.

Beehner, J. C., Nguyen, N., Wango, E. O., Alberts, S. C. & Altmann, J. 2006. The endocrinology of pregnancy and fetal loss in wild baboons. *Hormones and Behavior*, **49**, 688–699.

Blouin, M. S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution*, **18**, 503–511.

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

Carpenter, P. J., Pope, L. C., Greig, C., Dawson, D. A., Rogers, L. M., Erven, K., Wilson, G. J., Delahay, R. J., Cheeseman, C. L. & Burke, T. 2005. Mating system of the Eurasian badger, *Meles meles*, in a high density population. *Molecular Ecology*, **14**, 273–284.

van de Casteele, T., Galbusera, P. & Matthysen, E. 2001. A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology*, **10**, 1539–1549.

Castracane, V. D., Copeland, K. C., Reyes, P. & Kuehl, T. J. 1986. Pubertal endocrinology of yellow baboon (*Papio cynocephalus*): plasma testosterone, testis size, body weight, and crown-rump length in males. *American Journal of Primatology*, **11**, 263–270.

Csilléry, K., Johnson, T., Beraldi, D., Clutton-Brock, T., Coltman, D., Hansson, B., Spong, G. & Pemberton, J. M. 2006. Performance of marker-based relatedness estimates in natural populations of outbred vertebrates. *Genetics*, **173**, 2091–2101.

Cutrera, A. P., Lacey, E. A. & Busch, C. 2005. Genetic structure in a solitary rodent (*Ctenomys talarum*): implications for kinship and dispersal. *Molecular Ecology*, **14**, 2511–2523.

David, V. A., Sun, S., Zhang, Z., Shen, F., Zhang, G., Zhang, H., Xie, Z., Zhang, Y.-P., Ryder, O. A., Ellis, S., Wildt, D. E., Zhang, A. & O'Brien, S. J. 2006. Parentage assessment among captive giant pandas in China. In: *Giant Pandas: Biology, Veterinary Medicine and Management* (Ed. by D. E. Wildt, A. Zhang, H. Zhang, D. L. Janssen & S. Ellis), pp. 245–273. Cambridge: Cambridge University Press.

Dittus, W. P. J. 1988. Group fission among wild toque macaques as a consequence of female resource competition and environmental stress. *Animal Behaviour*, **36**, 1626–1645.

Fernández, J. & Toro, M. A. 2006. A new method to estimate relatedness from molecular markers. *Molecular Ecology*, **15**, 1657–1667.

Fernández, J., Villanueva, B., Pong-Wong, R. & Toro, M. Á. 2005. Efficiency of the use of pedigree and molecular marker information in conservation programs. *Genetics*, **170**, 1313–1321.

Garant, D. & Kruuk, L. E. B. 2005. How to use molecular marker data to measure evolutionary parameters in wild populations. *Molecular Ecology*, **14**, 1843–1859.

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

Guo, S. W. 1996. Variation in genetic identity among relatives. *Human Heredity*, **46**, 61–70.

Guo, S.-W. & Xiong, M. 1997. Genes preserved in relatives. *Human Heredity*, **47**, 138–154.

Hamilton, W. D. 1963. The evolution of altruistic behavior. *American Naturalist*, **97**, 354–356.

Jones, K. L., Glenn, T. C., Lacy, R. C., Pierce, J. R., Unruh, N., Mirande, C. M. & Chavez-Ramirez, F. 2002. Refining the whooping crane studbook by incorporating microsatellite DNA and leg-banding analyses. *Conservation Biology*, **16**, 789–799.

Kalinowski, S. T., Wagner, A. P. & Taper, M. L. 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, **6**, 576–579.

Kappeler, P. M., Wimmer, B., Zinner, D. & Tautz, D. 2002. The hidden matrilineal structure of a solitary lemur: implications for primate social evolution. *Proceedings of the Royal Society of London, Series B*, **269**, 1755–1763.

Kleven, O., Jacobsen, F., Robertson, R. J. & Lifjeld, J. T. 2005. Extrapair mating between relatives in the barn swallow: a role for kin selection? *Biology Letters*, **1**, 389–392.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., Brotherstone, S., Guinness, F. & Clutton-Brock, T. 2002. Antler size in red deer: heritability and selection but no evolution. *Evolution*, **56**, 1683–1695.

Legendre, P., Lapointe, F.-J. & Casgrain, P. 1994. Modeling brain evolution from behavior: a permutational approach. *Evolution*, **48**, 1487–1499.

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

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

Milligan, B. G. 2003. Maximum-likelihood estimation of relatedness. *Genetics*, **163**, 1153–1167.

Parrott, M. L., Ward, S. J. & Temple-Smith, P. D. 2007. Olfactory cues, genetic relatedness and female mate choice in the agile antechinus (*Antechinus agilis*). *Behavioral Ecology and Sociobiology*, **61**, 1075–1079.

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

Randall, D. A., Pollinger, J. P., Wayne, R. K., Tallents, L. A., Johnson, P. J. & Macdonald, D. W. 2007. Inbreeding is reduced by female-biased dispersal and mating behavior in Ethiopian wolves. *Behavioral Ecology*, **18**, 579–589.

Russello, M. A. & Amato, G. 2004. *Ex situ* population management in the absence of pedigree information. *Molecular Ecology*, **13**, 2829–2840.

Taylor, A. C., Horsup, A., Johnson, C. N., Sunnucks, P. & Sherwin, B. 1997. Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat *Lasiorhinus krefftii*. *Molecular Ecology*, **6**, 9–19.

Thomas, S. C. 2005. The estimation of genetic relationships using molecular markers and their efficiency in estimating heritability in natural populations. *Philosophical Transactions of the Royal Society of London, Series B*, **360**, 1457–1467.

Thomas, S. C., Coltman, D. W. & Pemberton, J. M. 2002. The use of marker-based relationship information to estimate the heritability of body weight in a natural population: a cautionary tale. *Journal of Evolutionary Biology*, **15**, 92–99.

Toro, M., Barragán, C., Óvilo, C., Rodrigañez, J., Rodriguez, C. & Silió, L. 2002. Estimation of coancestry in Iberian pigs using molecular markers. *Conservation Genetics*, **3**, 309–320.

Trontti, K., Aron, S. & Sundström, L. 2005. Inbreeding and kinship in the ant *Plagiolepis pygmaea*. *Molecular Ecology*, **14**, 2007–2015.

Van Horn, R. C., Buchan, J. C., Altmann, J. & Alberts, S. C. 2007. Divided destinies: group choice of female savannah baboons during social group fission. *Behavioral Ecology and Sociobiology*, **61**, 1823–1837.

Wade, T. D. 1979. Inbreeding, kin selection, and primate social evolution. *Primates*, **20**, 355–370.

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