



Microsatellites behaving badly: empirical evaluation of genotyping errors and subsequent impacts on population studies

A.C. Kelly¹, N.E. Mateus-Pinilla², M. Douglas², M. Douglas², P. Shelton³ and J. Novakofski¹

¹Department of Animal Sciences, University of Illinois, Urbana, IL, USA

²Illinois Natural History Survey, Institute of Natural Resource Sustainability, University of Illinois, Champaign, IL, USA

³Division of Wildlife Resources, Illinois Department of Natural Resources, One Natural Resources Way, Springfield, IL, USA

Corresponding author: A.C. Kelly

E-mail: amykelly@illinois.edu

Genet. Mol. Res. 10 (4): 2534-2553 (2011)

Received January 4, 2011

Accepted April 5, 2011

Published October 19, 2011

DOI <http://dx.doi.org/10.4238/2011.October.19.1>

ABSTRACT. Microsatellites are useful tools for ecological studies because they can be used to discern population structure, dispersal patterns and genetic relationships among individuals. However, they can also yield inaccurate genotypes that, in turn, bias results, promote biological misinterpretations, and create repercussions for population management and conservation programs. We used empirical data from a large-scale microsatellite DNA study of white-tailed deer (*Odocoileus virginianus*) to identify sources of genotyping error, evaluate corrective measures, and provide recommendations to prevent bias in population studies. We detected unreported mutations that led to erroneous genotypes in five of 13 previously evaluated microsatellites. Of the five problematic markers, two contained mutations that resulted in null alleles, and three contained mutations that resulted in imperfect repeats. These five microsatellites had error rates that were four times greater on average than those observed in the remaining eight.

Methodological corrections, such as primer redesign, reduced errors up to 5-fold in two problematic loci, although analytical corrections (computational adjustment for errors) were unable to fully prevent bias and, consequently, measures of genetic differentiation and kinship were negatively impacted. Our results demonstrate the importance of error evaluation during all stages of population studies, and emphasize the need to standardize procedures for microsatellite analyses. This study facilitates the application of microsatellite technology in population studies by examining common sources of genotyping error, identifying unreported problems with microsatellites, and offering solutions to prevent error and bias in population studies.

Key words: Microsatellite; Null alleles; Parentage; White-tailed deer; Genotyping; Heterozygote deficiency

INTRODUCTION

Microsatellites have become increasingly popular as genetic tools for the study of ecology, kinship and evolution (Chapuis and Estoup, 2007). These short tandem repeat DNA markers are extremely polymorphic and versatile, and when applied appropriately, provide a precise and statistically powerful means of quantifying animal movements, kinship, and genetic structure (Dewoody et al., 2006). Moreover, recent innovations have greatly reduced labor and costs for microsatellite analysis, thus promoting it as a cost-effective source of data for studies of dispersal and population structure.

Numerous ecological studies have used microsatellites to examine parentage, effective population size (McLean et al., 2008), gene flow (Epps et al., 2005), and relatedness (Lunn et al., 2000), and these data clearly elucidate population processes important for population management and conservation. In bighorn sheep (*Ovis canadensis*), microsatellite genotypes revealed a decrease in genetic diversity among California populations separated by interstate highways and canals, and supported over- and underpass construction to re-establish connectivity (Epps et al., 2005). Similarly, Coulon et al. (2006) were able to define “ecologically meaningful management units” for roe deer (*Capreolus capreolus*) in France based on genetic populations constrained by landscape barriers.

While microsatellites are clearly beneficial and their applicability widespread, few studies fully recognize the technical shortcomings that can arise when they are employed as a measure of genetic exchange (Dakin and Avise, 2004). Errors occur when the observed genotype is not equivalent to the true genotype, a result most often attributable to null alleles that fail to amplify in PCR assays because of mutations within microsatellite primer annealing sites. These mutations can reduce PCR efficiency and result in non-amplifying microsatellite alleles that appear to defy Mendelian inheritance patterns (Callen et al., 1993). Errors can also be caused by imperfect repeats, or by mutations within the microsatellite sequence that yield inconsistencies in the repeat motif (Figure 1). Imperfect repeats can be identical in size to perfect (non-mutated) repeats, which can result in incorrect scoring of two alleles as one. Furthermore, technical problems associated with PCR can create errors in microsatellite genotypes when smaller alleles are preferentially detected over larger alleles because of the greater

efficiency of amplifying shorter PCR fragments. These inconsistencies, often referred to as 'allelic dropout', tend to be associated with low DNA template quality, which is commonly reported in studies employing non-invasive tissue sampling (Taberlet et al., 1999).

Unrecognized errors, regardless of origin, can have adverse consequences in ecological studies. For example, Gagneux et al. (2001) were forced to retract a high-profile journal article (Gagneux et al., 1997) because the actual source of father-offspring mismatches in their data (previously interpreted as extra-group copulations based on paternal exclusion of males) was instead the result of undetected null alleles. This example demonstrates the broad and pervasive potential for biological misinterpretations that can result from genotyping errors, and underscores the importance of obtaining accurate baseline molecular information.

A substantial frequency of genotyping error is most often manifested in a population as homozygote excess (heterozygote deficiency). This occurs when heterozygous genotypes fail to amplify because of mutations or allelic dropout, and thus appear monomorphic. Usually, homozygote excess can be detected through deviations from Hardy-Weinberg equilibrium (HWE), but in some instances it only becomes apparent when confirmed parent-offspring dyads seemingly demonstrate different genotypes resulting from shared non-amplifying alleles (Dakin and Avise, 2004). Several software programs are designed to aid in the detection of genotyping errors by testing specifically for the presence of null alleles and allelic drop-out or by identifying heterozygote deficiencies via HWE tests (i.e., MicroChecker: Van Oosterhout et al., 2004; and GenAEx: Peakall and Smouse, 2006).

Heterozygote deficiencies, once identified, can sometimes be corrected to prevent bias during genetic analyses. Correction techniques fall into two broad classes: analytical or methodological. The former rely upon computational algorithms that quantify and subsequently rectify heterozygote deficiency by adjusting genetic data accordingly. GenePop (Rousset, 2008), for example, can estimate the overall frequency of null alleles based on heterozygote deficiencies, while MicroChecker (Van Oosterhout et al., 2004) can correct raw genotypes and account for the presence of non-amplifying alleles. Other corrective measures have been recommended by Wagner et al. (2006), who proposed adjusting standard relatedness statistics to account for the possibility of null alleles. Programs like Cervus adjust for errors by incorporating a user-specified error rate in parentage simulations (Kalinowski et al., 2007). However, many of these analytical corrections have limited utility for downstream analysis of genetic patterns. MicroChecker, for example, provides a list of corrected genotypes based on the estimated frequency of missing alleles (Van Oosterhout et al., 2004). Adjusted genotypes are not assigned to specific individuals; hence they cannot be used to analyze individual multi-locus genotypes. Also, most software packages fail to accept allele frequency data, thus preventing analyses based on frequency estimates of missing alleles (Dewoody et al., 2006).

Methodological approaches, on the other hand, often involve alternative primers or adjusted PCR conditions to boost the amplification of undetected alleles. This can require extensive laboratory effort and may not correct all sources of genotyping error, thus limiting information gain for the effort invested. An alternative is the amplification of additional polymorphic loci, followed by subsequent rejection of those that amplify inconsistently or deviate from HWE. This approach can be costly and time-consuming, particularly if multiple markers have non-amplifying alleles. Moreover, this approach is relatively inefficient in the long run because it does not engage the underlying molecular issues *per se*.

Despite shortcomings, both analytical and methodological approaches can be employed to correct for the effects of genotyping errors. Yet, remarkably, 90% of studies reporting null alleles failed to compensate for this error source and researchers instead elected to employ the affected loci 'as is' (Dakin and Avise, 2004). This is problematic, in that genotyping errors negatively impact analyses of genetic variability (Wagner et al., 2006), with the severity of bias depending on the type of analysis performed. For example, F-statistics are particularly sensitive because non-amplifying alleles can deflate within-population variance and exaggerate genetic differentiation (Chapuis and Estoup, 2007). Simulations have shown that in parentage studies, average exclusion probabilities are reasonably immune to the effects of null alleles, although their presence can significantly elevate the false exclusion rate for certain individuals (Dakin and Avise, 2004).

Sources of genotyping error and methods to estimate error rate (Bonin et al., 2004) have been thoroughly addressed, mostly with simulation studies and theoretical approaches (Paetkau, 2003). Several studies have focused on noninvasive sampling because it is especially prone to genotyping error, so much so that several noninvasive genotyping protocols have been designed specifically for error reduction (Woods et al., 1999; Paetkau, 2003). However, genotyping errors in large-scale studies with robust tissue samples and large DNA yields are less commonly addressed. Further, empirical reports of errors and their resulting effects on biological interpretations are relatively rare. Yet, evaluating microsatellite errors is critical, because consequences for biological misinterpretation could be severe, particularly when microsatellite data are used to guide conservation, management, or disease detection.

We utilized microsatellite markers previously optimized for use in white-tailed deer (*Odocoileus virginianus*) (Anderson et al., 2002) to demonstrate the effects of genotyping errors on population genetic studies and the benefits that can be gained from their correction. These microsatellites are of particular interest because many have been used in prior genetic studies of hunter-harvested species, with the intent to guide wildlife and infectious disease management (DeYoung et al., 2003a; Blanchong et al., 2008; Gear et al., 2010). We applied empirical data to 1) describe genotyping errors found during large-scale genetic analysis and dissect their molecular basis, 2) compare methodological and analytical techniques most frequently used to rectify genotyping errors, 3) determine how null alleles, imperfect repeats and allelic dropout can impact biological interpretations, and 4) provide recommendations on how genotyping errors can be minimized in population studies.

MATERIAL AND METHODS

Laboratory procedures

For genetic analyses, we used white-tailed deer (N = 877 total multi-locus genotypes, including 90 replicate and 65 fetus genotypes) sampled through Illinois Department of Natural Resources (IDNR) Chronic Wasting Disease (CWD) surveillance in north-central Illinois for genetic analyses. Previous genetic analyses conducted in this area revealed extensive deer movement and extremely low genetic structure, suggesting a single large population (Kelly et al., 2010).

We stored skeletal muscle or lymph node tissues of sampled animals in 100% ethanol or in a -80°C freezer. Genomic DNA was extracted using the Wizard Genomic DNA Purifica-

tion Kit (Promega, Madison, WI, USA), in accordance with manufacturer instructions. We genotyped individuals using 13 microsatellite primers that were optimized for use in white-tailed deer (Anderson et al., 2002), but were originally developed for cattle (*Bos taurus*), sheep (*Ovis aries*), goats (*Capra hircus*), and caribou (*Rangifer tarandus*). This panel included markers BM848, BM1225, BM4107, BM6506, CSN3 (Bishop et al., 1994), BM4208 (Talbot et al., 1996), IGF-1 (Kirkpatrick, 1992), OBCAM (Fries et al., 1993), OarFcb304 (Buchanan et al., 1993), RT20, RT23, RT27 (Wilson et al., 1997), and Srcrsp-10 (Bhebhe et al., 1994). We labeled forward primers with fluorescent dyes (NED, HEX, FAM) and separated microsatellite fragments on an ABI 3730XL capillary sequencer (Applied Biosystems, Foster City, CA, USA). We visualized microsatellite genotypes with GeneMapper, v. 4.0 (Applied Biosystems).

Detection of genotyping errors

We used MicroChecker, v. 2.2.3 (Van Oosterhout et al., 2004) to test for genotyping errors, and applied GenePop, v. 4.0 (Rousset, 2008) to examine deviations from HWE. We calculated observed (H_O) and expected (H_E) heterozygosities using 10,000 dememorization steps and 10,000 iterations in each of 100 batch runs, with significance ($P < 0.01$) based on the Markov chain algorithm (Rousset, 2008). To evaluate the stability of genotypes across independent PCRs, we genotyped 86 individuals in replicates of two or three, using all markers. Further, to verify Mendelian inheritance patterns, we genotyped 35 mother-offspring pairs (female deer with twins or triplets *in utero*; fetuses, $N = 65$). This allowed us to discover alleles that consistently failed to amplify, resulting in mismatched genotypes between mother and offspring. We counted mismatches among replicates and mother-offspring pairs as the number of genotypes with at least one erroneous allele. To calculate error rates, we divided the number of mismatches by the total number of genotypes for each locus.

We classified error sources into five exclusive categories: null alleles, imperfect repeats, miscalls, transcription errors, and allelic dropout. We sequenced microsatellites (methods described below) to confirm null alleles and imperfect repeats and compared microsatellite chromatograms with scored genotypes to identify errors in the other three categories. We classified mismatches between mother-offspring pairs and replicate samples as null alleles when microsatellites carried mutations in primer binding sites. We classified mismatches between mother-offspring pairs and replicate samples as imperfect repeats when microsatellites carried mutations in the repeat sequence. We identified miscalls as genotypes scored incorrectly upon first visual inspection (usually when aberrant stutter patterns or spurious peaks were present). Typing errors resulted from transcription errors during manual data editing, and allelic dropout (i.e., genotypes that scored heterozygous and homozygous in replicate reactions) resulted from inconsistent PCR amplifications.

Sample size effects

To determine if sample size affected our ability to detect HWE deviations, we randomly selected subsets of 25, 50, 100, 200, and 400 individuals for analysis and comparison with the full data-set of 722 genotyped deer (i.e., 877 total multi-locus genotypes with 65 fetuses and 90 replicates removed). For each subset, we calculated the inbreeding coefficient,

F_{IS} , and evaluated significance ($P < 0.01$) with GenePop, v. 4.0, for all loci, using 10,000 dememorization steps and 10,000 iterations in each of 100 batch runs.

Methodological corrections

We sequenced entire microsatellite loci, including primer binding sites and repeat sequences of markers CSN3, BM848, BM4208, BM6506, and RT20. We chose to sequence these five markers because they severely deviated from HWE, demonstrated inconsistent amplification, or showed mismatches between mother-offspring pairs (we will subsequently refer to these problems as major sources of error). Ten individuals were chosen for sequencing: five demonstrating mismatches or inconsistent amplifications and five that appeared to amplify normally. To evaluate sequences and ensure that they were of high quality, we used the Codon-Code Aligner v. 3.5.1 software (CodonCode Corporation, Dedham, MA, USA). Upon detection of mutations in microsatellite flanking regions of CSN3 and RT20, we relocated primer binding sites to areas with no genetic variability among sequenced samples: CSN3 (original reverse primer CSN3-R: GCACTTTATAAGCACCACAGC; redesigned reverse primer CSN3-R_R: TAGCTCATAATGTAAACCACTTT) and locus RT20 (original forward primer RT20-F: GCAGAAGAGTGAGTGTGAGT; redesigned RT20-F_R forward primer: TGG AAGATTTTCAGAAATGAT).

All individuals ($N = 877$) were then genotyped with redesigned primers RT20_R and CSN3_R. We compared genotypes from redesigned primers with raw genotypes to determine n_0 , the number of heterozygous genotypes incorrectly scored as homozygous. To calculate empirical estimates of null allele frequencies (r), we divided n_0 by the total number of alleles at each locus.

Analytical corrections

MicroChecker and GenePop refer to estimates of all non-amplifying alleles (regardless of their origin) as 'null allele frequency estimates'. For our purposes, this term will subsequently refer to calculated estimates of non-amplifying alleles resulting from null alleles, allelic dropout, or imperfect repeats. We compared four algorithms for estimating r across the five loci that deviated from HWE. First, we used MicroChecker to calculate r with analytical corrections described by Van Oosterhout et al., 2004 (AC-O), Chakraborty et al., 1992 (AC-C), Brookfield, 1996 (Equation 1; AC-B). We also employed GenePop to calculate null allele frequencies according to analytical methods described by Dempster et al., 1977 (AC-D). Based on deviations from HWE, we then applied MicroChecker to correct raw genotype frequencies by identifying the number of homozygous genotypes likely to be heterozygous for a non-amplifying allele. For locus BM848, we compared corrected genotypes (obtained from MicroChecker) to original genotypes to determine n_0 that was incorrectly scored as homozygous. Because MicroChecker does not perform genotype corrections for markers with null allele frequencies $< 5\%$, we could not use it to calculate n_0 for loci BM4208 and BM6506.

Evaluation of corrective measures

To determine if analytical and methodological corrections affected the outcome of

population analyses, we compared results from five different data sets: DS-O (the original 13 locus panel with null alleles or imperfect repeats), DS-C (the original panel with corrected frequency estimates for null alleles), DS-R (original panel with redesigned primers substituted for null alleles), DS-I (original panel with imperfect repeats removed), and DS-RI (original panel with imperfect repeats removed and redesigned primers substituted for null alleles) (Table 1). These five data sets allowed us to compare results based on equal numbers of loci and varying frequencies of error.

Table 1. Microsatellite data sets (from white-tailed deer, *Odocoileus virginianus*) used for genetic analysis of genotyping errors.

Microsatellite, description	Data set ^a				
	DS-O	DS-C	DS-R	DS-I	DS-RI
BM1225, no major errors detected	√	√	√	√	√
BM4107, no major errors detected	√	√	√	√	√
BM4208, imperfect repeat	√	√ ^b	√		
BM6506, imperfect repeat	√	√ ^b	√		
BM848, imperfect repeat	√	√ ^b	√		
CSN3, null allele	√	√ ^b		√	
CSN3 _R , redesigned primer			√		√
IGF-1, no major errors detected	√	√	√	√	√
OarFcb304, no major errors detected	√	√	√	√	√
OBCAM, no major errors detected	√	√	√	√	√
RT20, null allele	√	√ ^b		√	
RT20 _R , redesigned primer			√		√
RT23, no major errors detected	√	√	√	√	√
RT27, no major errors detected	√	√	√	√	√
Srersp-10, no major errors detected	√	√	√	√	√
Total	13	13	13	10	10
<i>r</i>	25%	0%	15%	10%	0%

^aMicrosatellite panels used to generate data sets were: DS-O = original panel; DS-C = panel with corrections based on null frequency estimates; DS-R = panel with redesigned primers; DS-I = original panel with imperfect repeats removed; DS-RI = panel with redesigned primers added and imperfect repeats removed. ^bIndicates use of corrected frequency as calculated per Brookfield et al., 1996; Equation 1. A √ indicates inclusion in panel. Total = number of markers in panel. *r* = panel-specific estimate of null allele frequencies (empirical *r* used for RT20, CSN3, *r*^{AC-B} used for BM848, BM6506, and BM4208).

When population analyses required corrected frequencies, we employed the methods of Brookfield, 1996 (Equation 1) for null allele frequency estimation (as recommended by Van Oosterhout et al., 2004). To determine *r* for each panel, we summed empirical estimates of null alleles for RT20 and CSN3 and calculated estimates for BM848, BM4208, and BM6506.

Corrective measures and population analyses

To determine the response of each marker to corrective measures, we individually evaluated problematic markers BM4208, BM6506, BM848, RT20, and CSN3. We subdivided the study area into 13 study sites (≥ 20 deer/site) and then calculated pairwise F_{ST} values for all study site combinations using allele frequencies from original genotypes, corrected frequencies, and frequencies generated with redesigned primers. To determine if these markers affected levels of genetic differentiation, we performed Fisher exact tests in Arlequin, v. 3.1 (Excoffier et al., 2005) using 100,000 Markov chain steps with 10,000 dememorization steps at $P < 0.05$. We summed the total number of significant tests for each data set and the num-

ber of pairwise tests that gained or lost significance. This allowed us to compare changes in genetic differentiation that resulted from methodological and analytical correction methods.

To ensure that differences in genetic differentiation were not solely attributable to variable explanatory power between data sets, we evaluated each locus using POWSIM (Ryman and Palm, 2006) to estimate the power to detect differentiation. For power analyses, we performed Fisher exact tests on allele frequencies using 1000 dememorization steps followed by 1000 iterations in each of 100 batches. We calculated power as the number of significant tests per 1000 replicates.

We examined the effects of genotyping errors on population-based distance measures and spatial autocorrelation with the Spatial Genetic v. 1.0 Software (Degen et al., 2001). This program derives calculations from allele frequencies rather than individual genotypes, thus permitting an evaluation of analytical correction methods using multi-locus data. For this analysis, we further subdivided the study area into 26 study sites (with approximately equal numbers of deer in each), so that we could evaluate effects of genotyping error on population analyses conducted at greater spatial resolution (~2 km²). We calculated pairwise F_{ST} 's and Nei's genetic distances among populations for all five data sets, and computed pairwise geographic distances from the centroid of each study site based on x- and y-coordinates. To test for spatial autocorrelation, we calculated average genetic distances within nine distance classes and compared these to average genetic distances estimated across all classes. We chose equidistant intervals, ensured that each distance class had at least 30 data points, and evaluated significance using 500 random permutations of the observed data (Degen et al., 2001).

Corrective measure and parentage analyses

To determine the effects of genotyping errors on parentage studies, we assigned parentage to 65 deer fetuses with Cervus, v. 3.0 (Kalinowski et al., 2007). Using 35 confirmed mothers and 35 randomly selected females as candidate mothers, we assigned parentage to fetuses and examined changes in parentage non-exclusion probabilities (i.e., the probability of excluding a non-parent) in the presence of genotyping errors. We based parentage assignments on data from four (DS-O, DS-R, DS-I, and DS-RI) of the five marker panels because Cervus cannot infer parentage from allele frequency data (i.e., DS-C). To determine the maximum false exclusion rate, we ran initial simulations with an assumed genotyping error rate of 0. However, in subsequent simulations, we adjusted the proportion of loci mistyped to account for panel-specific error rates (as recommended by Kalinowski et al., 2007). For each panel, we determined the false exclusion rate by dividing the number of unassigned fetuses by the total number of mother-fetus pairs. We calculated average non-exclusion probabilities within each simulation and for all loci combined (Kalinowski et al., 2007).

RESULTS

Genotyping errors

Of the original 13 microsatellites that we evaluated, six deviated from HWE (BM4208, BM6506, BM848, RT20, RT27, and RT23) and one amplified inconsistently (CSN3) when the original panel was evaluated. Of the six markers deviating from HWE, RT23 and RT27 did not

appear to be affected by null alleles, imperfect repeats, or allelic dropout, whereas sequencing confirmed a null allele in RT20 that resulted from a point mutation and a duplication in the 5' primer binding site. Microsatellites BM848, BM4208 and BM6506 were imperfect repeats, and we observed zero (perfect repeat), one (heterozygous) and two (homozygous) copies of point mutations within repeat regions of sequenced individuals. In addition CSN3, which did not deviate from HWE but amplified inconsistently, contained two point mutations in the 3' primer binding site that resulted in a null allele (Figure 1). Although loci RT23, and RT27 deviated from HWE in the full data set, we did not observe mismatches in mother-offspring pairs, and transcription error accounted for the single mismatch detected in these two markers (Table 2). Further, F_{IS} values were extremely low (Table 3), and HWE deviations in RT23 and RT27 were not apparent in data subsets or in larger data sets (data not shown), most likely suggesting subtle subpopulation structure rather than non-amplifying alleles. Null alleles and imperfect repeats were detected in primers originally designed for cattle (BM848, BM4208, BM6506, and CSN3) and caribou (RT20).

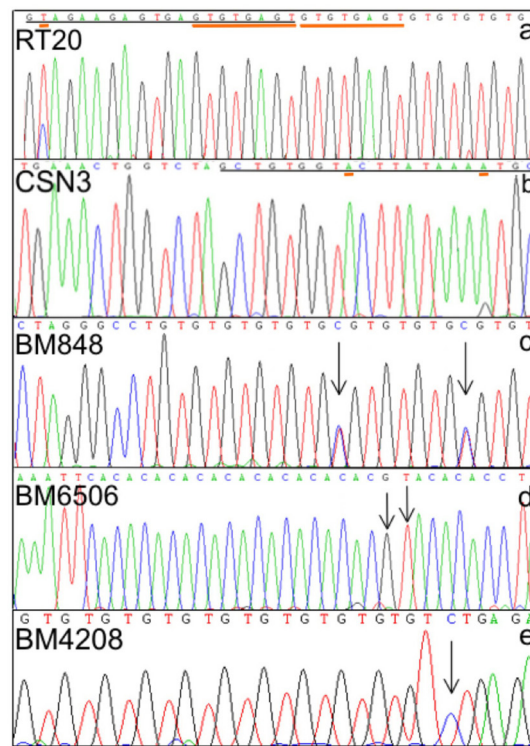


Figure 1. Molecular basis for null alleles and imperfect repeats detected in white-tailed deer microsatellites. **a.** RT20 with original forward primer underlined in black (RT20 original forward primer: GCAGAAGAGTGAGTGTGAGT) and a point mutation and duplication in primer binding site underlined in orange. **b.** CSN3 with original reverse primer binding site underlined in black (original reverse primer: GCACTTTATAAGCACCACAGC; reverse complement shown in chromatogram GCTGTGGTGCTTATAAAGTGC) with two base pair substitutions underlined in orange. **c.** BM848 imperfect repeat with multiple heterozygous base pair substitutions (concurrent red and blue peaks) indicated by arrows. **d.** BM6506 imperfect repeat with 'GT' insertion indicated by arrows. **e.** BM4208 imperfect repeat with a 'C' insertion indicated by arrow.

Table 2. Mismatch errors caused by null alleles in known pedigree and mother-fetus and replicate genotypes of white-tailed deer.

Marker	Mismatch mother-fetus ^a	Mismatch replicate ^b	Error rate
Srersp-10	0	0	0
CSN3 _R	0	0	0
CSN3	0	4	1.6
IGF-1	0	1	0.4
RT27	0	1	0.4
RT23	0	1	0.4
BM1225	0	1	0.4
OarFeb304	0	2	0.7
BM4107	0	2	0.7
OBCAM	0	3	1.1
BM848	3	0	1.2
BM6506	3	2	1.9
BM4208	3	5	3.1
RT20	7	4	4.5
RT20 _R	0	2	0.8

^aMismatched genotypes between mother-fetus pairs for 65 fetuses and 35 mothers. ^bMismatched genotypes for replicate samples from 86 individuals. Error rates were calculated by taking the total number of mismatched genotypes (mother-fetus and replicate) divided by the total number of genotypes * 100.

Table 3. Effects of sample size on null allele detection.

Marker	N = 25		N = 50		N = 100		N = 200		N = 400		N = Full (722)	
	F_{IS}	P	F_{IS}	P	F_{IS}	P	F_{IS}	P	F_{IS}	P	F_{IS}	P
BM848	0.07	ns	0.05	ns	0.01	ns	0.07	**	0.10	***	0.11	***
BM1225	-0.02	ns	0.02	ns	0.02	ns	-0.02	ns	0.00	ns	0.01	ns
BM4107	0.24	*	0.12	ns	0.04	ns	0.05	ns	0.04	ns	0.03	ns
BM4208	0.20	**	0.13	**	0.06	*	0.07	***	0.07	***	0.07	***
BM6506	0.00	ns	-0.04	ns	-0.01	ns	0.03	ns	0.03	*	0.02	*
CSN3	0.07	ns	-0.08	ns	-0.10	ns	-0.13	ns	0.01	ns	0.04	ns
CSN3 _R	-0.03	ns	-0.16	ns	-0.19	ns	-0.15	ns	-0.03	ns	0.03	ns
IGF-1	0.22	*	0.13	ns	0.08	ns	0.04	ns	0.03	ns	0.01	ns
OarFeb304	-0.12	ns	-0.09	ns	0.04	ns	0.01	ns	-0.01	ns	0.03	ns
OBCAM	-0.05	ns	0.01	ns	-0.01	ns	-0.05	ns	-0.04	ns	-0.01	ns
RT20	0.27	**	0.27	***	0.28	***	0.26	***	0.28	***	0.29	***
RT20 _R	0.00	ns	-0.02	ns	0.01	ns	-0.02	ns	-0.02	ns	-0.02	ns
RT27	0.04	ns	0.01	ns	0.04	ns	-0.01	ns	0.00	ns	0.01	*
RT23	-0.10	ns	0.00	ns	0.00	ns	-0.02	ns	0.00	ns	0.01	***
Srersp-10	0.05	ns	0.06	ns	0.02	ns	-0.04	ns	0.00	ns	0.01	ns

Subsets of 25, 50, 100, 200, and 400 individuals were randomly selected for analysis and compared with the full data-set (722 deer). ns = $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Tests for heterozygote deficiency were performed and F_{IS} values were evaluated for significance using 10,000 dememorization steps followed by 10,000 iterations in each of 100 batches. Markers highlighted in gray demonstrated changes in significance for heterozygote deficiency at different values of N.

Among the five data sets derived from different marker panels, overall frequency of non-amplifying alleles (r : calculated empirically for RT20 and CSN3 and estimated by Brookfield, 1996; Equation 1 for BM848, BM4208, and BM6506) ranged from 25% with the original panel (DS-O) to 0% in the original panel with redesigned primers added and imperfect repeats removed (DS-RI) (Table 1). Nearly two-thirds of all genotyping errors resulted from null alleles and imperfect repeats, although transcription errors also contributed substantially

to mismatched genotypes (Figure 2). Overall error rates for individual markers ranged from 0 (Srcrsp-10 and CSN3_R) to 4.5% (RT20), with an average error rate of 1.1% (Table 2).

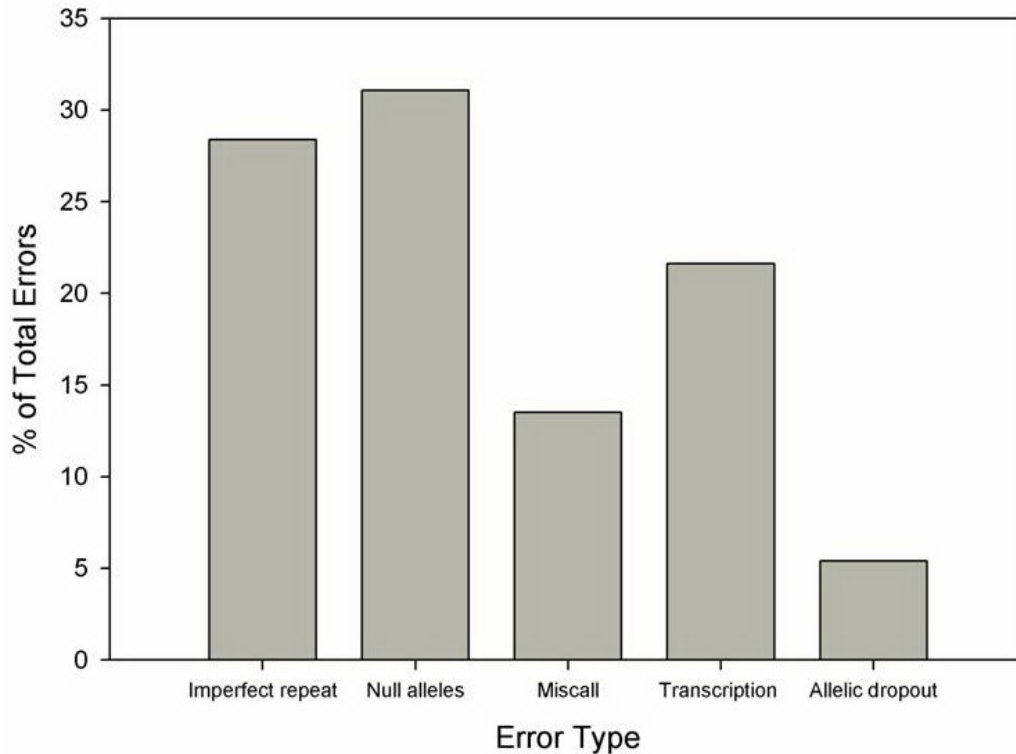


Figure 2. Types of genotyping errors detected in genetic characterization of mother-offspring pairs (65 fetuses and 35 mothers) and 86 individuals genotyped in replicate. The % of total errors was calculated as ((number of errors for each category / number of errors observed for all categories) * 100).

Sample size effects

The ability to detect HWE deviations was profoundly influenced by the sample size of the data set (Table 3). Eight of 15 markers (13 original + two redesigned) experienced changes in significance level changes (two lost and six gained significance) when varying numbers of individuals were included in the analysis. For markers BM848, BM6506 and RT20, F_{IS} and its associated significance increased as individuals were added to the data set. For marker BM6506, 400 individuals had to be analyzed before a significant departure from HWE was observed. In contrast, F_{IS} and its significance decreased for BM4107 and IGF-1 as sample size increased. Although significance also changed for markers RT27 and RT23, F_{IS} values were generally low; unstable and significant departures from HWE were detected only in the full data set.

Methodological corrections

After identifying null alleles in RT20 and CSN3, we redesigned primers to bind to segments of DNA that did not vary among sequenced individuals. These redesigned primers consistently amplified alleles, with resulting genotypes in HWE. Mutations within the repeat sequence cannot be corrected using methodological techniques, and they were thus not applied to imperfect repeats of BM848, BM4208, and BM6506. Although we were unable to correct for errors associated with imperfect repeats in BM4208, we were able to eliminate spurious bands in this marker by decreasing the number of PCR cycles from 25 to 20. This reduction prevented self-annealing of PCR products, thereby increasing the detectability of microsatellite alleles (Bovo et al., 1999).

Analytical corrections

The four analytical methods to estimate null allele frequency produced identical results for BM4208 ($r = 0.04$) and BM6506 ($r = 0.01$) (Table 4). Estimates based on AC-O, AC-B, and AC-D were identical ($r = 0.05$) for BM848, while values estimated by AC-C were slightly higher ($r = 0.06$). Empirical null allele estimates (determined from genotypes that went from homozygous to heterozygous with redesign primers) were 0.02 for CSN3 and 0.13 for RT20. In comparison, methods AC-O and AC-C correctly estimated the frequency of null alleles for CSN3, while methods AC-B and AC-D produced slight underestimates ($r = 0.01$). For RT20, only AC-B accurately estimated the null allele frequency, while the other three methods produced overestimates. Based on these results, the method AC-B (Brookfield, 1996) produced the most accurate estimates of null alleles in our study.

Table 4. Descriptive statistics for white-tailed deer microsatellites with null alleles before and after primer redesign.

Marker	Original primers					Redesigned primers			Null allele estimates				
	N	n_0	H_o	H_e	P	H_o	H_e	P	Empirical	r^{AC-C}	r^{AC-B}	r^{AC-D}	r^{AC-O}
RT20	716	192	0.62	0.87	*	0.87	0.86	ns	0.13	0.17	0.13	0.14	0.14
CSN3	710	31	0.48	0.50	ns	0.49	0.50	ns	0.02	0.02	0.01	0.01	0.02
BM848	716	65	0.76	0.85	***	na	na	na	na	0.06	0.05	0.05	0.05
BM4208	716	nc	0.83	0.90	***	na	na	na	na	0.04	0.04	0.04	0.04
BM6506	716	nc	0.86	0.88	ns	na	na	na	na	0.01	0.01	0.01	0.01

N = number of genotypes; n_0 = number of genotypes that contained null or non-amplifying alleles; H_o = observed heterozygosity; H_e = expected heterozygosity calculated by GenePop and evaluated for significance using the Markov chain algorithm; nc = not calculable; na = not applicable; ns = $P > 0.05$; * $P < 0.05$; *** $P < 0.001$. Empirical null allele estimates (r) were calculated by taking the number of null alleles (determined from genotypes that went from homozygous to heterozygous with redesign) divided by the total number of alleles analyzed. r^{AC-C} = null allele estimate according to Chakraborty et al. (1992); r^{AC-B} = null allele estimate according to Brookfield (1996); r^{AC-D} = null allele frequency according to Dempster et al. (1977), and r^{AC-O} = null allele frequency according to Van Oosterhout et al. (2004).

Corrective measures and population analyses

For all five markers tested, average pairwise F_{ST} values were higher when calculated with null alleles. This overestimation was so severe for RT20 that the distribution of F_{ST} values

calculated with null alleles was substantially higher (one-tailed F value = 2.9; $P = 0.057$) than the distribution calculated with corrected and redesigned frequencies. The use of corrected null allele frequencies appeared to rectify this problem, as average F_{ST} values for corrected and redesigned data sets were nearly identical. Variable results were observed for all loci examined when the number of different study sites was used as a metric for genetic differentiation (Figure 3). Genetic differentiation was higher for BM4208 and RT20 when errors were included as compared to corrected and redesigned frequencies (Figure 3). In contrast, markers BM6506, CSN3 and BM848 showed higher levels of genetic differentiation with corrected frequencies when compared to original frequencies including null alleles and imperfect repeats (Figure 3). The differences in genetic differentiation are attributable to errors rather than variations in genetic resolution, as changes in power averaged less than 2% among data sets for the same locus.

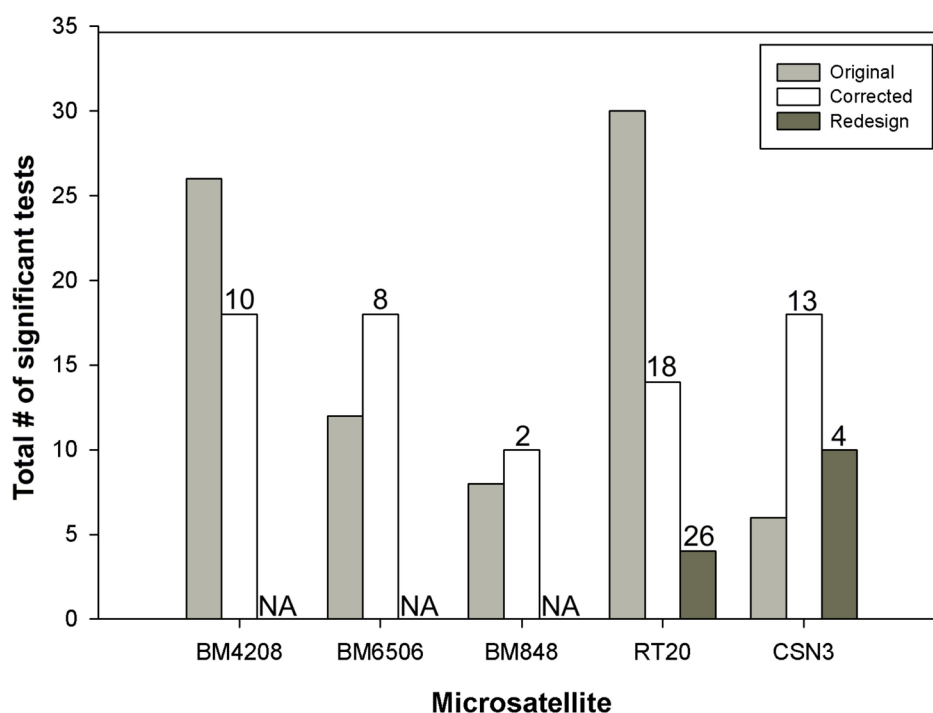


Figure 3. Effects of genotyping error on estimation of genetic differentiation. Fisher exact tests for genetic differentiation were performed using original, corrected and redesigned allele frequencies of individual microsatellites with 100,000 Markov chain steps and 10,000 dememorization steps in Arlequin. The number of significant tests at $P < 0.05$ was summed and used as a metric for overall genetic differentiation. The number of pairwise comparisons that changed in significance (gain or loss compared to original primer) are shown above the bars. NA = not applicable; primers were not redesigned for microsatellites with imperfect repeats.

Although population differentiation was sensitive to genotyping errors, population-based measures and spatial autocorrelation were only slightly affected. Overall patterns of spatial autocorrelation were not affected by genotyping error as none of the data sets showed consistent patterns of spatial structure to begin with. However, average F_{ST} values and Nei's genetic distances were variable for data sets with equal numbers of loci. When 13 loci were

employed, data set DS-C yielded the highest distance values, followed by DS-O and DS-R, although differences were not significant. For data sets based on 10 loci, genetic distances were higher, (although not significantly so) for DS-I than for DS-RI.

Corrective measures and parentage analyses

Parentage assignments for known mother (35) and fetus (65) pairs revealed that false exclusion rates increased dramatically as the frequency of null alleles increased (Figure 4). The original data set ($r = 25\%$) had the highest false exclusion rate (20%), with 13 of 65 confirmed mothers being excluded because of mismatches with their fetuses. On the other hand, the original data set with redesigned primers added and imperfect repeats removed ($r = 0\%$) did not falsely exclude any confirmed mothers. There was also an inverse relationship between the estimated frequency of null alleles within a data set and the probability of non-exclusion. The original data set with redesigned primers added and imperfect repeats removed had the highest probability of non-exclusion ($P = 0.004$), while the original data set had a much lower probability of non-exclusion ($P = 0.0003$). However, these changes were mostly attributable to higher resolving power and not solely to errors, as data sets with equal numbers of loci (original vs original with redesigned, original minus imperfect repeats vs original with redesigned minus imperfect repeats) had similar estimates of probabilities of non-exclusion (<1% change in P).

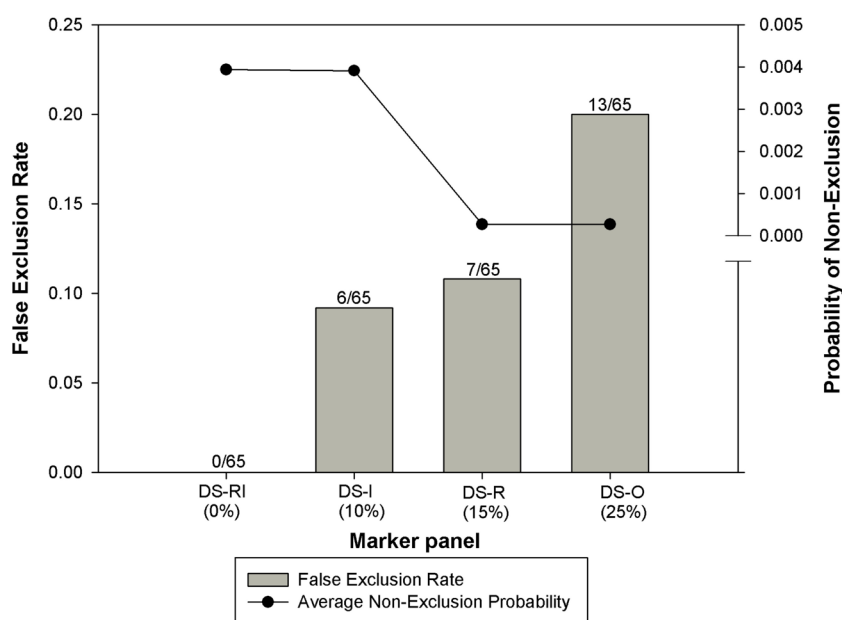


Figure 4. Effects of genotyping errors on parentage analysis. False exclusion rates and probabilities of non-exclusion for original (DS-O), original with redesigned primers (DS-R), original with imperfect repeats removed (DS-I), and original with redesigned primers added and imperfect repeats removed (DS-RI). Panel specific estimates of null allele frequency (r) are shown below each panel name. Parentage was assigned to 65 deer fetuses using 70 candidate mothers (35 known mothers and 35 random females). The proportion of individuals falsely excluded due to mismatches from null or non-amplifying alleles is shown above each bar. Probabilities of non-exclusion were averaged across all loci and reported for each panel tested.

Panel-specific error rates (calculated from mismatches in replicate samples and from mother-fetus pairs) were specified for parentage analyses in Cervus, although they failed to rectify problems with false exclusion for genotypes based on the original data set. The proportion of mistyped loci was 1.3% for DS-O, yet 13 of 65 confirmed mothers (20%) were still excluded. On the other hand, when panel specific error rates were used (0.5% for DS-RI, 1% for DS-I and 1.1% for DS-R), genotypes based on the other three data sets did not falsely exclude any confirmed mothers.

DISCUSSION

Microsatellite markers have truly impacted population studies by providing a wealth of genetic information for conservation biology, molecular ecology and population genetics. These disciplines, among many others, have benefited greatly because microsatellite analysis provides an efficient way of estimating complex population parameters like dispersal, kinship, inbreeding, and population size. However, the ability of microsatellites to elucidate such intricate genetic patterns depends greatly on the accuracy of genotypes.

Genotyping errors

Using empirical data, we were able to demonstrate that null alleles and imperfect repeats accounted for the vast majority of genotyping errors in our study. The loci that we employed are generally considered to be valid markers (Anderson et al., 2002) and have been used extensively in population studies (DeYoung et al., 2003b; Blanchong et al., 2008; Gear et al., 2010). The fact that >30% of our original panel contained confirmed null alleles or imperfect repeats suggests that more stringent criteria for marker evaluation should be considered. The validation study of Anderson et al. (2002) included markers BM848, BM6506, and BM4208, and while H_o was $< H_e$ for all three, the authors conclude that “little evidence exists of null alleles at these loci”, with a recommendation for use in studies of gene flow and parentage (Anderson et al., 2002). Moreover, markers BM4208 and BM6505 continue to be employed in studies of kinship and relatedness (Gear et al., 2010), despite demonstrated HWE deviations (Anderson et al., 2002; Comer et al., 2005) and inconsistent genotypes (DeYoung et al., 2003a) in deer populations across the US. To prevent data bias and biological misinterpretations, markers need to be assessed thoroughly with each novel application. Furthermore, to increase awareness of problematic markers in the scientific community, genotyping errors should be reported, especially when deviations from HWE are detected (Dewoody et al., 2006).

Sample size effects

Elevated sample sizes greatly enhanced our ability to detect low frequency null alleles, imperfect repeats, and effects of sample size on HWE deviations. Sample size issues are a consistent problem for ecological studies (Rao, 2001), especially those involving rare or endangered species (Storfer, 1996). Our results suggest that elevated sample sizes are necessary to detect markers with null allele frequencies $\leq 5\%$. BM848 required $N = 200$ and BM6506 required $N = 400$ before HWE deviations became significant. The differences in F_{IS} that we observed at varying sample sizes were not likely attributable to population substructuring

because the sampled population is not genetically structured at this spatial scale (Kelly et al., 2010). Rather, our results reflect an inability of small sample sizes to capture the range of genetic variation in our population. This is contrary to suggestions of Takezaki and Nei (1996) and DeYoung et al. (2003b) that 50 individuals are sufficient for estimating population allele frequencies. While 50 may suffice for genetically homogeneous populations, our deer population was highly mobile and genetically admixed (Kelly et al., 2010). Simulation studies have supported these results by demonstrating that sample sizes <100 are unlikely to produce reliable estimates of population allele frequencies (Rao, 2001). Moreover, HWE inconsistencies observed during subsampling show that dividing a large population into smaller groups for analysis of HWE (as suggested in MicroChecker for $N > 750$; Van Oosterhout et al., 2004) can lead to erroneous conclusions about population-level deviations.

Methodological corrections

We found that sequencing the primer-binding sites of suspect microsatellites was an effective way to resolve the molecular basis for HWE deviations and amplification inconsistencies. Further, information about the molecular origins of errors can be useful for determining which type of corrective measures can best rectify the problem. For example, duplications in the RT20 primer binding site produced multiple fragments corresponding to the same allele, and corrected genotypes based on estimated null frequencies were unable to compensate for the complexity of this issue. Redesigning the primer, on the other hand, resulted in HWE and a drastic decrease in genetic differentiation as compared to full and corrected frequencies. Additionally, confirmation that a microsatellite contains an imperfect repeat would prevent time and effort wasted in a redesign of primers, for this will not resolve the issue. By understanding the molecular basis of genotyping errors and the methods used to estimate their frequencies, one can *a priori* predict the utility of correction methods. In general, the cost of sequencing a few individuals is much lower than the cost of testing additional markers, and this is an excellent method for exploring errors, and gaining the necessary information for targeted re-genotyping.

Corrective measures and population analyses

Null alleles, imperfect repeats, and allelic dropout can contribute to an overestimation of observable alleles, a decrease in observed heterozygosity, and an increase in the apparent level of inbreeding (Dewoody et al., 2006). These impacts bias allele frequencies and promote an overestimation of genetic differentiation (Chapuis and Estoup, 2007), as observed in this study for markers RT20 and BM4208. For CSN3, however, genetic differentiation was diminished for those genotypes with null alleles, as compared to genotypes derived from redesigned primers (Figure 3). Relocating primer binding sites promoted amplification of novel alleles at this locus, which, in turn, resulted in greater differentiation. Because CSN3 is not very polymorphic ($H_E = 0.5$) and null allele frequencies were low (2%), an increase in discriminatory power from additional alleles was greater than the diversity lost because of null alleles. Still, variance in genetic differentiation observed among all of our markers would have led to alternate conclusions about the intensity of gene flow among study sites. Collectively these findings suggest that the influences of genotyping errors on population differentiation are variable, depending upon genetic diversity of the marker and the frequency of the null allele.

Despite their impacts on parentage and population differentiation, null alleles and imperfect repeats did not affect overall patterns of spatial autocorrelation. Nevertheless, we observed differences in Nei's genetic distances and F_{ST} values for panels with and without these errors. As noted in other studies, analyses of population structure tend to be less sensitive to errors than individual-based analyses (Taberlet et al., 1999), and in our study this was most likely because bias inflicted by erroneous microsatellites was diminished by the rest of the markers in the panel. Multi-locus analysis is more realistic than our single-locus analysis of genetic differentiation, because most population studies apply panels that are likely to contain combinations of markers with and without null alleles or imperfect repeats. The fact that spatial autocorrelation results were qualitatively similar suggests that null allele frequencies $\leq 25\%$ will have minimal effects on coarse patterns of population structure, although differences in distance measures will undoubtedly be affected even at frequencies as low as 10%.

Corrective measures and parentage analyses

Consistent with simulation studies (Dakin and Avise, 2004), false exclusion rates in parentage analyses increased dramatically as the estimated frequency of null (or non-amplifying) alleles increased. This problem can be circumvented by only excluding candidate parents if mismatches occur at more than one locus. Our results suggest this *post hoc* approach may be too conservative for preventing false exclusions, as we detected null alleles and imperfect repeats in almost 40% of the markers employed, and most importantly, we detected mismatches between mother-fetus pairs across multiple loci. An alternative correction would be to specify a user-defined error rate to account for genotyping errors during parentage assignment, thus reducing the number of false exclusions (Kalinowski et al., 2007). This improved false exclusions in data sets with null allele estimates of $\leq 15\%$, but it failed when estimates were 25%. This is most likely because null alleles violate Kalinowski et al. (2007) assumptions of random erroneous genotypes and equal error rates across all loci, but could also be due to the underestimation of error rates in this study. When null allele frequencies are low ($< 20\%$), their effects on parentage are often minimal (Dakin and Avise, 2004). While this conclusion was confirmed in our study (at least with regards to exclusion probabilities), we still noted that low frequency null alleles ($< 15\%$) yielded appreciable false exclusion probabilities. In general, null alleles can hinder parentage analyses and corrective measures do not always remedy their negative effects.

Collectively, our findings emphasize the need for consistent evaluation of genotyping error throughout all stages of population studies. Researchers cannot rely solely on validation and simulation studies to provide recommendations for certain microsatellite markers. The utility of a marker panel is dependent on the research objectives at hand, and the effects of genotyping errors are far too stochastic to allow for generalized statements about their pervasiveness or the appropriateness of corrective measures.

There are several ways that researchers can prevent genotyping errors and subsequent bias in biological interpretations of genetic data. Some have suggested choosing a panel of candidate microsatellites that was developed specifically for the target species or a phylogenetically close relative, as decreasing phylogenetic distance generally decreases the probability of null alleles (Chapuis and Estoup, 2007) and imperfect repeats. However, our data warrant caution as markers designed for cattle (*B. taurus*; suborder Ruminantia, family Bovidae; BM848, BM4208, BM6506) had lower frequencies of null alleles in white-tailed deer than

did markers designed for caribou (*R. tarandus*; suborder Ruminantia, family Cervidae; RT20), even though the latter belong to the same taxonomic family.

Despite these cautions, one can still employ a marker with a null allele if primers are relocated to a binding region that is invariant across the target population. This approach is especially useful because microsatellites are not widely available for all organisms and some studies must rely on primers designed for non-target species. Regardless, researchers should prepare for potential non-amplification, monomorphism, or imperfect repeats by testing a surplus of candidate markers. By doing so, poor-performing markers can be eliminated early on and resources allocated towards optimizing others that exhibit a higher probability of success.

When optimizing microsatellites, markers should be tested on individuals from different populations or genetic cohorts, as this will enhance the potential for determining polymorphism and detecting null alleles or imperfect repeats. Species mobility and population admixture will undoubtedly influence the distribution of alleles, and these must be considered when estimating the number of individuals required to establish baseline allele frequencies. Contrary to what the literature suggests, a reliable estimate of genetic parameters for highly admixed populations may require more than 50 individuals.

Many studies report an increase in genotyping errors as DNA quantity and quality decrease (Taberlet et al., 1999). DNA extracted from feces, hair (Taberlet et al., 1996), and museum specimens (Horváth et al., 2005) can be extremely difficult to genotype and can often produce unreliable and error-prone genotypes. Our own experience suggests that formalin fixed or improperly preserved tissues had extremely high failures and error rates, and DNA from these types of samples consistently yield inferior template quality. In addition, protocols should be in place to ensure reliable and repeatable amplification for all individuals. The assumption that human errors are the primary cause of non-amplification could lead to the inclusion of null alleles in the data set.

Our findings reiterate the need to evaluate errors for the entire duration of the genotyping process. Software packages like MicroChecker (Van Oosterhout et al., 2004) should be continually employed to check for HWE and non-amplifying alleles, even as additional samples are genotyped. We recommend independent replication of at least 5-10% of samples for assessment of marker-specific error rates, and when possible the inclusion of samples with known pedigrees can greatly facilitate the identification of invisible null alleles. If null alleles are suspected, an estimate of their frequency can help determine the utility of the marker for various genetic analyses, but it should not be used exclusively to predict the level of bias that the null alleles will impart.

Our study provides empirical insight into the unpredictable nature of genotyping errors and their effects on measures of population structure and parentage. We detected several null alleles and imperfect repeats in microsatellites that were developed and oft-employed in managed species, which introduced substantial bias in our results. The inclusion of these errors would suggest that populations were genetically more distinct than they were in reality. Further, we would have been unable to determine paternity for a large proportion of samples, despite the fact that true parents were included. Biased conclusions such as these could lead to the false identification of management units, or inaccurate inferences about breeding systems.

Additionally, analytical approaches to correct for genotyping errors were unable to fully compensate for this bias. Caution should be exercised when these techniques are employed. Our results demonstrate that genotyping errors are study-specific, and additional re-

search is needed to develop standardized error detection protocols and further evaluate corrective measures. We also encourage researchers to publish error rates and deviations from HWE for each microsatellite so that problematic markers can be avoided in the future. Established standards for genotyping and increased error reporting would result in a higher level of accuracy in population studies, which would improve the overall quality of information available to the scientific community. Although the focus of this study was on microsatellites designed for white-tailed deer, our recommendations are generally applicable to all studies that employ these markers.

ACKNOWLEDGMENTS

The Illinois Department of Natural Resources collected all samples and we thank the numerous Wildlife Biologists who generously donated both time and effort to this study. The Illinois Department of Agriculture Animal Disease Laboratory (Galesburg) also provided invaluable assistance.

REFERENCES

- Anderson JD, Honeycutt RL, Gonzales RA, Gee KL, et al. (2002). Development of microsatellite DNA markers for the automated genetic characterization of white-tailed deer populations. *J. Wildl. Manag.* 66: 67-74.
- Bhebe E, Kogi J, Holder DA, Arevalo E, et al. (1994). Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-6, SR-CRSP-7, SR-CRSP-8, SR-CRSP-9 and SR-CRSP-10 loci. *Anim. Genet.* 25: 203.
- Bishop MD, Kappes SM, Keele JW, Stone RT, et al. (1994). A genetic linkage map for cattle. *Genetics* 136: 619-639.
- Blanchong JA, Samuel MD, Scribner KT, Weckworth BV, et al. (2008). Landscape genetics and the spatial distribution of chronic wasting disease. *Biol. Lett.* 4: 130-133.
- Bonin A, Bellemain E, Bronken EP, Pompanon F, et al. (2004). How to track and assess genotyping errors in population genetics studies. *Mol. Ecol.* 13: 3261-3273.
- Bovo D, Ruge M and Shiao YH (1999). Origin of spurious multiple bands in the amplification of microsatellite sequences. *Mol. Pathol.* 52: 50-51.
- Brookfield JF (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* 5: 453-455.
- Buchanan FC, Littlejohn RP, Galloway SM and Crawford AM (1993). Microsatellites and associated repetitive elements in the sheep genome. *Mamm. Genome* 4: 258-264.
- Callen DF, Thompson AD, Shen Y, Phillips HA, et al. (1993). Incidence and origin of "null" alleles in the (AC)_n microsatellite markers. *Am. J. Hum. Genet.* 52: 922-927.
- Chakraborty R, De Andrade M, Daiger SP and Budowle B (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Ann. Hum. Genet.* 56: 45-57.
- Chapuis MP and Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* 24: 621-631.
- Comer CE, Kilgo JC, D'Angelo GJ, Glenn TC, et al. (2005). Fine-scale genetic structure and social organization in female white-tailed deer. *J. Wildl. Manag.* 69: 332-344.
- Coulon A, Guillot G, Cosson JF, Angibault JMA, et al. (2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Mol. Ecol.* 15: 1169-1679.
- Dakin EE and Avise JC (2004). Microsatellite null alleles in parentage analysis. *Heredity* 93: 504-509.
- Degen B, Petit R and Kremer A (2001). SGS - Spatial Genetic Software: a computer program for analysis of spatial genetic and phenotypic structures of individuals and populations. *J. Hered.* 92: 447-449.
- Dempster AP, Laird NM and Rubin DB (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.* 39: 1-38.
- Dewoody J, Nason JD and Hipkins VD (2006). Mitigating scoring errors in microsatellite data from wild populations. *Mol. Ecol. Notes* 6: 951-957.
- DeYoung RW, Demarais S, Honeycutt RL, Gonzales RA, et al. (2003a). Evaluation of a DNA microsatellite panel useful

- for genetic exclusion studies in white-tailed deer. *Wildl. Soc. Bull.* 31: 220-232.
- DeYoung RW, Demarais S, Honeycutt RL, Rooney AP, et al. (2003b). Genetic consequences of white-tailed deer (*Odocoileus virginianus*) restoration in Mississippi. *Mol. Ecol.* 12: 3237-3252.
- Epps CW, Palsboll PJ, Wehausen JD and Roderick GK (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecol. Lett.* 8: 1029-1038.
- Excoffier L, Laval G and Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.
- Fries R, Eggen A and Womack JE (1993). The bovine genome map. *Mamm. Genome* 4: 405-428.
- Gagneux P, Woodruff DS and Boesch C (1997). Furtive mating in female chimpanzees. *Nature* 387: 358-359.
- Gagneux P, Woodruff DS and Boesch C (2001). Retraction: Furtive mating in female chimpanzees. *Nature* 414-508.
- Grear DA, Samuel MD, Scribner KT, Weckworth BV, et al. (2010). Influence of genetic relatedness and spatial proximity on chronic wasting disease infection among female white-tailed deer. *J. Appl. Ecol.* 47: 532-540.
- Horváth MB, Martínez-Cruz B, Negro JJ, Kalmár L, et al. (2005). An overlooked DNA source for non-invasive genetic analysis in birds. *J. Avian Biol.* 36: 84-88.
- Kalinowski ST, Taper ML and Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16: 1099-1106.
- Kelly AC, Mateus-Pinilla NE, Douglas M, Brown W, et al. (2010). Utilizing disease surveillance to examine gene flow and dispersal in white-tailed deer. *J. Appl. Ecol.* 47: 1189-1198.
- Kirkpatrick BW (1992). Identification of a conserved microsatellite site in the porcine and bovine insulin-like growth factor-I gene 5' flank. *Anim. Genet.* 23: 543-548.
- Lunn NJ, Paetkau D, Calvert W, Atkinson S, et al. (2000). Cub adoption by polar bears (*Ursus maritimus*): determining relatedness with microsatellite markers. *J. Zool.* 251: 23-30.
- McLean JE, Seamons TR, Dauer MB, Bentzen P, et al. (2008). Variation in reproductive success and effective number of breeders in a hatchery population of steelhead trout (*Oncorhynchus mykiss*): examination by microsatellite-based parentage analysis. *Conserv. Genet.* 9: 295-304.
- Paetkau D (2003). An empirical exploration of data quality in DNA-based population inventories. *Mol. Ecol.* 12: 1375-1387.
- Peakall R and Smouse PE (2006). Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6: 288-295.
- Rao C (2001). Sample size considerations in genetic polymorphism studies. *Hum. Hered.* 52: 191-200.
- Rousset FCO (2008). GenePop'007: a complete re-implementation of the GenePop software for Windows and Linux. *Mol. Ecol. Res.* 8: 103-106.
- Ryman N and Palm S (2006). POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Mol. Ecol. Notes* 6: 600-602.
- Storfer A (1996). Quantitative genetics: a promising approach for the assessment of genetic variation in endangered species. *Trends Ecol. Evol.* 11: 343-348.
- Taberlet P, Griffin S, Goossens B, Questiau S, et al. (1996). Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res.* 24: 3189-3194.
- Taberlet P, Waits LP and Luikart G (1999). Noninvasive genetic sampling: look before you leap. *Trends Ecol. Evol.* 14: 323-327.
- Talbot J, Haigh J and Plante Y (1996). A parentage evaluation test in North American elk (Wapiti) using microsatellites of ovine and bovine origin. *Anim. Genet.* 27: 117-119.
- Takezaki N and Nei M (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144: 389-399.
- Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4: 535-538.
- Wagner AP, Creel S and Kalinowski ST (2006). Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity* 97: 336-345.
- Wilson GA, Strobeck C, Wu L and Coffin JW (1997). Characterization of microsatellite loci in caribou Rangifer tarandus, and their use in other artiodactyls. *Mol. Ecol.* 6: 697-699.
- Woods JG, Paetkau D, Lewis D, McLellan BN, et al. (1999). Genetic tagging of free-ranging black and brown bears. *Wildl. Soc. Bull.* 27: 616-627.