Received: 8 March 2021 | Revised: 5 July 2021 | Accepted: 8 July 2021

DOI: 10.1111/1755-0998.13466

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Pedigree simulations reveal that maternity assignment is reliable in populations with conspecific brood parasitism, incomplete parental sampling and kin structure

Caroline M. Thow[1] | John M. Eadie[2] | Caitlin P. Wells[2,3] | Bruce E. Lyon[1]

[1]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California, USA

[2]Department of Wildlife, Fish, and Conservation Biology, University of California, Davis, California, USA

[3]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, Colorado, USA

Correspondence
Bruce E. Lyon, Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA.
Email: belyon@ucsc.edu

Funding information
National Science Foundation, Grant/Award Number: 1354894 and 1355208; D. G. Raveling Endowment; USDA National Institute of Food and Agriculture, Grant/Award Number: 1013603

Abstract

Modern genetic parentage methods reveal that alternative reproductive strategies are common in both males and females. Under ideal conditions, genetic methods accurately connect the parents to offspring produced by extra-pair matings or conspecific brood parasitism. However, some breeding systems and sampling scenarios present significant complications for accurate parentage assignment. We used simulated genetic pedigrees to assess the reliability of parentage assignment for a series of challenging sampling regimes that reflect realistic conditions for many brood-parasitic birds: absence of genetic samples from sires, absence of samples from brood parasites and female kin-structured populations. Using 18 microsatellite markers and empirical allele frequencies from two populations of a conspecific brood parasite, the wood duck (*Aix sponsa*), we simulated brood parasitism and determined maternity using two widely used programs, CERVUS and COLONY. Errors in assignment were generally modest for most sampling scenarios but differed by program: CERVUS suffered from false assignment of parasitic offspring, whereas COLONY sometimes failed to assign offspring to their known mothers. Notably, COLONY was able to accurately infer unsampled parents. Reducing the number of markers (nine loci rather than 18) caused the assignment error to slightly worsen with COLONY but balloon with CERVUS. One potential error with important biological implications was rare in all cases—few nesting females were incorrectly excluded as the mother of their own offspring, an error that could falsely indicate brood parasitism. We consider the implications of our findings for both a retrospective assessment of previous studies and suggestions for best practices for future studies.

KEYWORDS
assignment errors, brood parasitism simulations, conspecific brood parasitism, kinship, microsatellites, parentage assignment

1 | INTRODUCTION

Identifying the parents of offspring from a pool of candidate parents is critical for estimating the reproductive effort and success of individuals in populations, and for understanding the evolution of reproductive strategies generally. However, the existence of alternative reproductive tactics such as conspecific brood parasitism and extra-pair matings often makes assigning parentage and estimating

reproductive success difficult because the social parent of an offspring may not be its genetic parent (Griffith et al., 2002; Lyon & Eadie, 2017; Walling et al., 2010). To detect these alternative reproductive tactics, researchers have employed both observational and genetic methods. The widespread use of genetic techniques, in particular, has allowed researchers to identify otherwise cryptic parents and elucidate breeding dynamics that were previously undetected and therefore poorly understood (e.g., Andersson et al., 2019; Brouwer & Griffith, 2019). Molecular markers such as microsatellite DNA and single nucleotide polymorphisms (SNPs) can both exclude putative parents by detecting mismatches between offspring and their social parents, and identify the true parents of an offspring from a pool of candidate parents (Arnold & Owens, 2002; Avise et al., 2002; Jones et al., 2010; Petrie & Møller, 1991).

Conspecific brood parasitism (CBP) is an alternative reproductive tactic that is often difficult to detect without genetic techniques. Conspecific brood parasites lay eggs in the nest of other females of the same species without providing any further care for those offspring. This form of parasitism is widespread in birds and has now been reported in over 250 species (Yom-Tov, 2001; Yom-Tov & Geffen, 2017). Historically, CBP has been under-detected and often undetected (Yom-Tov, 2001), since it is difficult to detect parasitism with behavioural observations alone, and in many species it is impossible to correctly identify parasitic eggs based on visual cues alone (Eadie et al., 2010; Macwhirter, 1989; Yom-Tov, 1980). Depending on the species, a parasitic female in a given year may have her own nest in addition to laying parasitically, or the parasite may forgo nesting altogether (Lyon & Eadie, 2017)—parasites without their own nests usually escape detection unless they are caught in the act of parasitism. A clear understanding of which females lay parasitically, and why, has been limited by the ability to identify the parasites and the contexts in which they engage in parasitism (Lyon & Eadie, 2017).

The development of molecular markers has increased the frequency at which CBP has been detected (Arnold & Owens, 2002) and enabled researchers to determine which females in the population lay parasitic eggs, and under what circumstances (Lyon & Eadie, 2008). Although molecular techniques provide powerful tools for investigating CBP, they are not without error and several factors can contribute to the risk of misassignments (Kalinowski et al., 2007; Jones et al., 2010; Lemons et al., 2014). These include having too few markers, which leads to low exclusion probabilities (Harrison et al., 2013; Jones et al., 2010; Lemons et al., 2014) and a variety of genotyping errors, including allelic dropout or null alleles, which can lead to allelic mismatches and false exclusion of true parents (Hoffman & Amos, 2005).

Errors in assignment are particularly important in the study of CBP because in addition to adding noise to a study, they may incorrectly suggest that something interesting biologically has taken place when in fact it has not. There are two distinct types of assignment error: (i) incorrectly leaving offspring unassigned (*false exclusion* of the true mother), or (ii) incorrectly assigning offspring to the wrong mother (*false assignment* of the true mother's progeny to another female). These lead to different errors of inference about the biology of CBP, including nuances that depend on the context of the reproductive tactic (i.e., parasitic status) of the mother of the incorrectly assigned or unassigned offspring (Table 1). For example, when studies consider unassigned offspring as parasitic (e.g., Lesobre et al., 2010; Tucker et al., 2016) and assignment power is low (e.g., studies with a small number of markers), there is a risk that many of these unassigned "parasites" are actually nonparasitic offspring belonging to the female of the nest in which they are found. In kin-structured populations, another source of error might arise with important implications. Specifically, cases where offspring are incorrectly assigned to a close relative of the true mother could be interpreted as evidence of kin-directed parasitism when in fact there may be none. This is especially relevant for studies of waterfowl (order Anseriformes) for which a number of recent studies suggest that CBP may be kin-directed and cooperative rather than parasitic (Andersson et al., 2019; Jaatinen et al., 2011; Nielsen et al., 2006; Tiedemann et al., 2011); however, the degree to which assignment errors might be biased towards kin has not been fully explored.

Studies of waterfowl have played a leading role in understanding CBP, in part because parasitism is particularly common in this group (Yom-Tov, 1980) but also because they have an interesting constellation of reproductive attributes that make CBP potentially complex and intriguing. However, some of these same complexities may complicate the accurate detection and interpretation of CBP. For example, strong female natal philopatry in waterfowl creates kin-structured populations, and so parasites could lay eggs in the nest of relatives (Andersson, 1984; Eadie et al., 1988). Kin parasitism could make it more difficult to distinguish among candidate mothers (Double et al., 1997; Jaatinen et al., 2011; Olsen et al., 2001) and potentially result in incorrect assignment of offspring to a relative of the true mother (Jones et al., 2010). Compounding this issue, in many waterfowl species parasitic females often do not have nests of their own (Lyon & Eadie, 2017), and DNA samples for adult females are typically obtained only from nesting females because they can be captured at their nests (Jaatinen et al., 2009; Lemons & Sedinger, 2011; Nielsen et al., 2006). This could potentially bias assignments towards individuals present in the candidate parent pool (Araki & Blouin, 2005; R. Nielsen et al., 2001). Finally, in most waterfowl, males are not involved in nesting and samples from potential sires are often missing (Jaatinen et al., 2011; Tiedemann et al., 2011); when assignments are made based solely on the maternal half of parentage, exclusion may be more difficult (Double et al., 1997). In isolation, each of these challenges may have only a small impact on the accuracy of parentage assignment, but since they often occur jointly, their combined effects could be substantial.

A variety of approaches have examined the reliability of parentage assignment including: examining the effect of genotyping errors (e.g., allelic dropout, null alleles, stuttering) and incomplete parent sampling on assignment accuracy in natural populations (Araki & Blouin, 2005; Berger-Wolf et al., 2007; Hoffman & Amos, 2005), comparing the results from assignment programs to ecological data to verify the accuracy of assignment programs (Guerier et al., 2012; Sánchez-Tójar et al., 2015; Walling et al., 2010), and determining the

TABLE 1 The potential consequences of incorrect assignments and unassigned offspring based on the type of error, the relationship of the false mother to the true mother is applicable, and which nest the offspring is located in. Note we only discuss consequences for offspring of non-nesting parasites, since those are the only parasites included in our simulations. We do not cover the consequences of incorrect assignment of the offspring of nesting females since this was a relatively rare occurrence in our simulations, particularly when all loci were included in the analyses.

			Consequences			
True mother	Assigned mother	Result	Frequency of CBP	RS nester	RS parasite	Kinship
Nesting	True nesting mother	Correctly assigned	Accurate	Accurate	—	—
Nesting	Parasite nonrelative	Incorrectly assigned	Inflated	Underestimated	Inflated	—
Nesting	Parasite relative	Incorrectly assigned	Inflated	Underestimated	Inflated	Inflated
Nesting	None (mother in sample)	Incorrectly unassigned	Inflated[a]	Underestimated	Inflated[a]	—
Nesting	None (mother not in sample)	Correctly unassigned	Inflated[a]	Underestimated	Inflated[a]	—
Parasite	True parasite mother	Correctly assigned	Accurate	—	Accurate	—
Parasite	Host nonrelative	Incorrectly assigned	Underestimated	Inflated	Underestimated	—
Parasite	Host relative	Incorrectly assigned	Underestimated	Inflated	Underestimated	Underestimated
Parasite	Other nonrelative	Incorrectly assigned	Underestimated	—	Underestimated	—
Parasite	Other relative	Incorrectly assigned	Underestimated	—	Underestimated	Inflated[b]
Parasite	None (Mother in sample)	Incorrectly unassigned	Underestimated[c]	—	Underestimated	—
Parasite	None (mother not in sample)	Correctly unassigned	Underestimated[c]	—	Underestimated	—

Abbreviations: RS = reproductive success.

[a]If unassigned offspring are assumed to be parasitic, as some studies do.

[b]Kinship possibly inflated among parasite females.

[c]If unassigned parasite offspring are excluded from the analysis.

effect of number of loci on the accuracy of parentage assignment (Karaket & Poompuang, 2012; Walling et al., 2010). However, few studies have addressed the combined effects of incomplete parent sampling and kinship on assignment accuracy (e.g., Double et al., 1997). One powerful method to explore the intersection of incomplete sampling and kinship is through the use of pedigree simulations (Harrison et al., 2013; Jones et al., 2010).

In this study, we used simulated populations to evaluate the accuracy and reliability of parentage assignment under conditions frequently observed in studies of CBP. We created populations with completely known pedigrees based on observed allele frequency distributions from our study populations of wood ducks (*Aix sponsa*), a cavity-nesting species in which CBP is common (Nielsen et al., 2006; Odell & Eadie, 2010; Semel & Sherman, 2001). Simulating populations with CBP allows us to assess the accuracy of parentage assignments because true parentage is perfectly known (Harrison et al., 2013). Our simulations included mixtures of related and unrelated females, as well as a mixture of parasitic and nonparasitic offspring. With these simulations, we investigated which characteristics of the candidate parent pool most influenced accurate genetic parentage assignment by conducting a tiered series of simulations with increasingly challenging candidate parent pool characteristics relative to a simulation with complete information: (i) missing sires (addressed by removing male genotypes from the candidate parent pool); (ii) missing parasitic females, the group of particular interest to CBP (addressed by removing a subset of candidate parasite mother genotypes from the candidate parent pool); and (iii) relatedness between females (addressed by including kin structure in our simulated populations with full-sibling and half-sibling mothers). We assessed the influence of these factors using the two most commonly used parentage assignment software programs, CERVUS (Kalinowski et al., 2007) and COLONY (Jones & Wang, 2010), which further allowed us to determine if any assignment problems are program-specific (e.g., Harrison et al., 2013). Finally, since the number of markers is known to affect statistical power of parentage assignment, we explored how the number of markers intersects with the different sampling regimes by repeating analyses with half the total number of loci.

2 | MATERIALS AND METHODS

2.1 | The wood duck study system and sample collection

Wood ducks are cavity-nesting waterfowl that readily nest in nestboxes. Females typically lay 10–15 eggs in their own nest; previous studies indicate that parasitized clutches contain between 16 and 22 eggs on average (Bellrose & Holm, 1994), but they can have as many as 58 eggs in our study populations (Odell & Eadie, 2010). Females display natal philopatry and can nest near and/or parasitize nesting relatives. Previous studies indicate that wood ducks are probably socially monogamous for a given nest and potentially serially monogamous in multibrooded populations, but the actual mating system has yet to be accurately determined (Baldassarre, 2014; Bellrose & Holm, 1994). Males associate with females prior to nesting but do not provide parental care to their offspring.

We obtained the genotypes used in our simulations from wild wood duck populations in Yolo County, CA. From 2012 to 2015, we monitored 237 wood duck nestboxes at four sites: Conaway Ranch in Woodland, CA, Putah Creek and Russell Ranch in Davis, CA, and Roosevelt Ranch in Zamora, CA. We trapped females on their nests between the first and third week of incubation to band them and obtain blood samples. We made additional efforts during the breeding season to band and blood sample non-nesting females in nest traps. To capture previously unbanded female ducks we also deployed bait traps, designed to capture ducks away from the nest, after the breeding season. We were unable to sample many male wood ducks in our populations due to their reclusive nature during the breeding season, and so we derived simulated male genotypes using allele frequencies from the females in all populations combined (see below).

We collected blood samples (~0.2 ml) via medial metatarsal venipuncture using a 28-gauge needle. We then either applied the sample to DNA-preserving filter paper (Adventec MFS) or collected the blood with a capillary tube and deposited it in a 0.5-ml sample of Queen's lysis buffer (Seutin et al., 1991). We sent unprocessed blood and samples of previously extracted DNA—extracted using either DNEasy spin column kits (Qiagen) or plate extraction (Whitehead laboratory; Tables S1 and S2) —to the UC Davis Veterinary Genetics Laboratory (VGL) where they were genotyped using 18 microsatellite loci developed for waterfowl (Tables S1 and S2). All DNA profiles for each marker were read twice: automatically using STRAND software (Toonen & Hughes, 2001), and then manually by a trained analyst to ensure accuracy of genotype calls. Details of genotyping methods will be forthcoming in a separate paper (manuscript in preparation). The field-collected wood duck genotypes were then used to generate the simulated wood duck genotypes used in our analyses.

2.2 | Simulated populations

We investigated the accuracy of maternity assignments under several potentially challenging contexts: when paternal genotypes are missing, the female sample is incomplete, or related females are in the population. We explored this by simulating different breeding populations of wood ducks, each with a subset of non-nesting parasitic females that varied in their relatedness to nesting hosts. We simulated wood duck populations of known pedigree, relatedness and reproductive tactic to produce offspring that could then be assigned to their parents under the various scenarios. We created two simulated populations of wood ducks based on empirical allele frequencies from two wild populations, located at Conway and Roosevelt Ranches. We conducted two separate sets of full simulations to ensure replicability and because, while our simulated genotypes were not subject to allelic errors, differences in allele frequencies could influence our results (they did not; Figure S1). Constructing pedigrees with exact relatedness, simulating known parasitism and tracking parasitic vs. nonparasitic offspring of known parentage was complex and time-consuming; since we found few differences between these simulations, we felt that analysing the results for two separate populations under each set of scenarios was sufficient to account for population differences at the regional and local scales that we were most interested in.

To obtain each simulated population with the desired kin structure, three generations were required (Figure 1): each female and male in the first generation (G1) was created from actual allele frequencies sampled from our wild wood duck populations, a second generation (G2) of females of known relatedness and pedigree was produced from a simulated mating of the first generation, and a third generation (G3) of offspring was created from a simulated mating of the second-generation females with additional males drawn at random from the same pool used to create the first generation of males. The parentage assignments then involved matching the offspring from G3 with parents from G2.

To generate G1 simulated female genotypes we compiled two sets of field-sampled wood duck genotypes, one from each of two of our study sites: Conaway Ranch females ($N = 52$, collected between 2012 and 2015) and Roosevelt Ranch females ($N = 70$, collected in 2014 and 2015). We used additional field-sampled wood duck genotypes from two other study sites, Putah Creek females ($N = 10$) and Russell Ranch ($N = 10$), to generate G1 male genotypes. In wood ducks, females show strong natal philopatry but males do not and we took this into account so that our simulations would accurately reflect realistic population structure: we simulated female genotypes using allele frequencies from each of two real populations (Conaway Ranch, Roosevelt Ranch), whereas we simulated male genotypes using allele frequencies obtained from *all* of our populations (i.e., the full population-wide pool of possible alleles). For females, we used ML RELATE (Kalinowski et al., 2006) to determine the allele frequencies of each sample of females and used these frequencies to generate two independent sets of 200 simulated female genotypes with COLONY, one for each site. We similarly used COLONY to generate two sets of 136 simulated male genotypes (for the two population simulations), using allele frequencies from 40 randomly selected wild female genotypes (10 from each of our four real populations).

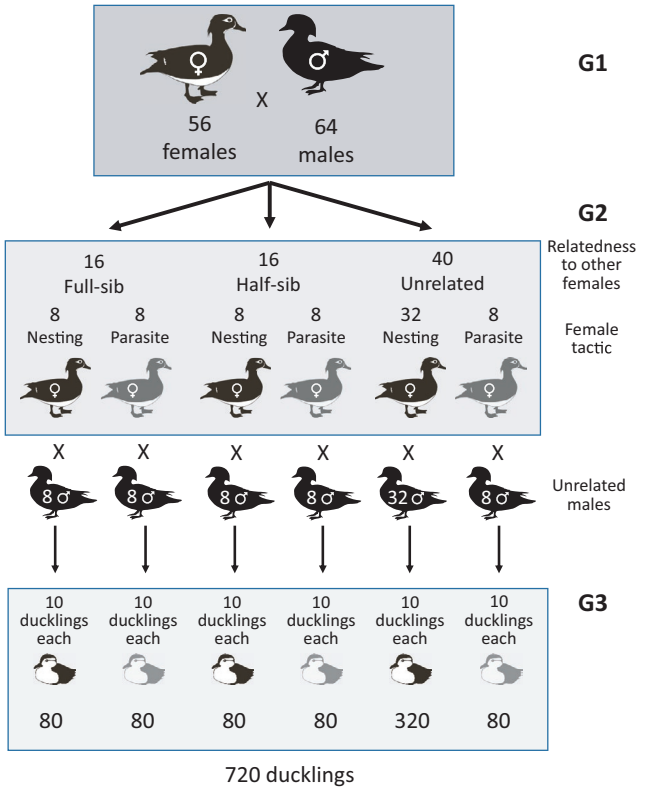

FIGURE 1 Overview of simulated genotypes used in this study, including the relatedness and nesting status of simulated second-generation (G2) females ($N = 72$) that were the mothers of the offspring used in the maternity assignments. The relatives were divided between the two breeding tactics: parasite and nester. There were eight pairs of full-siblings with one member in each of the two breeding categories; the same pattern applies to the half-siblings. The remaining females did not have relatives in the population, and were divided between nesters ($N = 32$) and parasites ($N = 8$). All females laid 10 eggs; the nesters laid their eggs in their own nest while the parasites did not have nests but laid all of their eggs in a nester's nest

For each of the two population simulations, we randomly chose 64 simulated males to pair with 56 G1 females to produce 72 G2 simulated female offspring. These numbers were chosen to produce the desired number of full- and half-sibling pairs of females (Figure 1). To produce each G2 female genotype, we randomly selected one allele from each G1 parent for each locus. Of the 72 G2 females, 16 were full-siblings, 16 were maternal half-siblings and 40 were not related to another female in the population (Figure 1). Full-sibling pairs were generated by having eight sets of G1 parents each produce two offspring. The half-siblings were generated by having eight G1 females have one offspring with each of two mates (hence the need for an additional eight extra males compared to females for the G1 individuals).

We then assigned G2 females to one of two reproductive tactics, resulting in 48 nesting females and 24 parasitic females. Sixteen of the nesting females had relatives that were brood parasites (Figure 1); the remaining 32 nesting females were not related to any other female in the population. The nesting females only reproduced through nesting, and the parasitic females only reproduced parasitically. In the field, identification of specific parasitic females depends on whether a genetic sample is obtained from the female. When parasites also have their own nests, researchers should be able to obtain samples from the parasites because birds are typically captured at their nests, and therefore are included as candidate parents. Parasites that do not have their own nest are more likely to go unsampled, and therefore are potentially at greater risk of having their offspring assigned incorrectly to another female because their genotypes are not included in the candidate parent pool. We restricted our simulations to parasitism by females without their own nests because we felt this presents the most challenging case for parentage assignments; the lack of samples from parasitic females could most affect the outcome of maternity analyses. Note, however, that it would be easy to use our nesting females to explore the ability to assign parasitic eggs by assuming that a fraction of the eggs they laid were in another female's nest. Assignment programs do not consider which nest an egg is laid in to assign maternity; whether or not an egg is deemed parasitic is determined by the researcher after the fact, by comparing maternity assignments with the nesting female (putative mother) for the nest each offspring is in. We can simulate nesting parasitism simply by arbitrarily assigning offspring to nests other than the nest of their mother and then determine how often they are detected as parasites (assigned to their true mother).

The primary goal for each simulation was to assign the maternity of the ducklings in the third generation to females in the second generation. To create the third-generation ducklings, we paired each of the 72 G2 females with a unique mate from the 72 unused male genotypes remaining in the male G1 pool to produce 10 ducklings per pair (720 ducklings in total). We constructed each duckling genotype by randomly selecting one allele from each parent for each locus. We then sought to identify the mothers of these 720 ducklings and, when errors in assignment were made, we tallied the specific type of biological error that resulted: (i) offspring left unassigned (*false exclusion* of the true mother), and (ii) offspring misassigned to the wrong mother (*false assignment* of the true mother's progeny to another female). A summary of each of the scenarios indicating the individuals that were included, presence of relatives, number of loci, number of candidate females and offspring generated under each simulation for two populations is provided in Table 2.

2.3 | Maternity analyses

We used COLONY and CERVUS to obtain maternity assignments for each set of simulated offspring. Both programs use maximum likelihood approaches to assign offspring to their parent(s), but they differ in a few key respects. CERVUS takes a pairwise maximum likelihood approach to assign offspring to their parents, using a three-step process. First, the program runs a pairwise parentage analysis on a simulated population created from the allele frequencies of the input genotypes from a population of interest, generating an LOD statistic

TABLE 2 Description and sample sizes of all scenarios indicating what individuals were included, presence or absence of relatives, number of loci, number of candidate females and offspring generated under each simulation for two populations

Scenario	Individuals included	Loci	Relatives	Number of candidate females	Number of offspring
All Parents	Includes all parents in the candidate parent pool, including paternal genotypes and the genotypes of all brood parasites as well as all nesting females in the population	All loci (18) Half loci (9)	Without relatives	48 × 2 populations 32 nesting, 16 parasitic	480 × 2 populations = 960
		All loci (18) Half loci (9)	With relatives	72 × 2 populations 48 nesting, 24 parasitic 16 fill-sib (FS), 16 half-sib (HS), 40 unrelated	720 × 2 populations = 1,440
All Females	Excludes all males, but includes both nesting and parasitic females	All loci (18) Half loci (9)	Without relatives	48 × 2 populations 32 nesting, 16 parasitic	480 × 2 populations = 960
		All loci (18) Half Loci (9)	With relatives	72 × 2 populations 48 nesting, 24 parasitic 16 FS, 16 HS, 40 unrelated	720 × 2 populations = 1,440
Nesting Parents	Includes only nesting females and their mates and excludes parasitic females and their mates	All loci (18) Half loci (9)	Without relatives	32 × 2 populations 32 nesting	480 × 2 populations (parasitic offspring of omitted parents included) = 960
		All loci (18) Half loci (9)	With relatives	48 × 2 populations 48 nesting 8 FS, 8 HS, 32 unrelated	720 × 2 populations (parasitic offspring of omitted parents included) = 1,440
Nesting Females	Includes only nesting females and excludes all males and parasitic females	All loci (18) Half loci (9)	Without relatives	32 × 2 populations 32 nesting	480 × 2 populations (parasitic offspring of omitted mothers included) = 960
		All loci (18) Half loci (9)	With relatives	48 × 2 populations 48 nesting 8 FS, 8 HS, 32 unrelated	720 × 2 populations (parasitic offspring of omitted mothers included) = 1440

(likelihood of the odds) score for each parent–offspring pair. Second, from this simulation, CERVUS identifies the critical LOD score, which is the LOD threshold value associated with a level of confidence. For example, the critical LOD score for a 95% confidence level is determined as the LOD score value above which 19 of 20 parents selected by the simulation as the most likely parents are the actual parents of the offspring. CERVUS allows the user to assign a strict and relaxed level of confidence, which are set by default at 95% and 80% confidence level thresholds respectively. While the authors of CERVUS recommend only accepting assignments made at the 95% or greater confidence level, some researchers have used assignments made at the 80% confidence level (Table 3). By definition, assignments made at less than the designated relaxed confidence level (usually 80%) indicate that the program cannot find a suitable parent in the candidate parent pool and the offspring is not assigned. In the final step, CERVUS uses the critical LOD scores generated by simulation to assign parents to offspring in the focal population: the candidate parent with the highest LOD score for an offspring is selected and a confidence level for the assignment is determined based on the value of that LOD score relative to the critical LOD scores for the predesignated confidence levels.

In contrast, COLONY takes a pedigree approach to assign parentage by determining the familial relationships best supported by the genetic evidence. COLONY uses an annealing algorithm to search for the best maximum likelihood pedigree configuration among thousands of possibilities to assign paternity, maternity and sibships (both full and half) in the population of interest. The user can specify details such as mating system, relatedness between known candidate parents, candidate parents to exclude for designated offspring, number of offspring per parent pair, number of runs and duration of each run. Unlike CERVUS, COLONY can also infer genotypes of individuals that were not sampled and include those genotypes as candidate parents to assign offspring to missing individuals. COLONY also differs in how it determines confidence in an assignment: it reports probabilities for each assignment, which it calculates from the proportion of pedigree configurations that included that assignment out of the total

number of pedigree configurations considered during a run. To be consistent with CERVUS, we considered any COLONY assignments with a probability of <.80 to be made at a low confidence level and we considered these offspring unassigned, as well as those COLONY assigned to inferred missing parents.

For each simulated wood duck population, we ran four separate analyses in both COLONY and CERVUS to assess the ability of the programs to assign offspring to their true mother when males are absent and when female sampling is incomplete. The configurations of the different analyses were chosen to explore the influence of different types of missing information, but each configuration also represents a sampling scenario that applies to actual breeding systems studied to date (e.g., Åhlund & Andersson, 2001; Eadie, 1989; Forslund & Larsson, 1995; McRae & Burke, 1996). These configurations vary in whether parasites and/or males are included in the genetic samples. The All Parents analysis served as a best-case scenario: we included all parents in the candidate parent pool, including paternal genotypes and the genotypes of all brood parasites as well as nesting females in the population. The All Females analysis excluded all males, but included both nesting and parasitic females as candidate parents. The Nesting Parents analysis included only nesting females and their mates as candidate parents, excluding parasitic females and their mates. Finally, the Nesting Females analysis included only nesting females and excluded males and parasitic females from the candidate parent pool. These four scenarios represent a gradient from complete sampling (All Parents), which is quite rare in most field studies, to the most limited sampling (Nesting Females), which is quite common given the logistic constraints of capturing all non-nesting females and males in many studies. As the reference set of scenarios, none of these analyses included relatives (maternal full- or half-siblings) in the candidate parental pool. We created populations without relatives by subsampling each of our original G2 populations to exclude one of each pair of related females. In these populations, all individuals without relatives were included (40 individuals, 32 nesting and eight non-nesting) but only one individual from each full-sibling pair was included (eight individuals of 16; these were non-nesting females) and no individuals from the half-sibling pairs were included (0 out of 16). This resulted in populations of 48 candidate mothers, 48 candidate fathers and 480 offspring. Of the 48 candidate mothers, 32 were nesting females and 16 were parasitic females.

2.4 | Effect of relatives and number of loci

We next compared the maternity analyses described above to analyses of the same four scenarios with relatives included. We preserved the 2:1 ratio of nesting to parasitic females present in the no-relative populations (48 nesting females, 24 parasitic females), adding in the full- and half-siblings generated from the G2 simulations. In these analyses, the 720 offspring were derived from three categories of mothers: those with no relatives in the population ($N = 40$), one half-sibling sister ($N = 16$) or one full-sibling sister ($N = 16$; Figure 1).

To explore how the number of loci might influence our analyses, we repeated all of the analyses described above twice: once with the full set of 18 loci, and again with a set of nine randomly selected loci (Tables S1 and S2). For the reduced loci analyses, we chose nine loci because it is half of the full set of loci and it is also the average number of loci used to detect CBP in 28 studies conducted across bird taxa in the past 12 years (Table 3).

2.5 | Program parameters used

For all analyses in both programs, we set the proportion of mothers or parents that were assumed to have been sampled to reflect the actual proportion of mothers or parents included in the current run (0.67 for runs where non-nesting parasitic females were excluded, 1.00 for runs where all candidate parents were included). We selected an accurate proportion of mothers to present the best-case scenario for the conditions we explored. Investigating variation in the proportion of parents sampled revealed that this parameter does not change the identity of parents assigned but instead affects the confidence level of assignment: overestimating the proportion of parents sampled inflates confidence of assignment, while underestimating the proportion of females sampled reduces confidence (Figure S2). For all COLONY analyses, we did not include any prior information on sibship among offspring or their parents and we allowed polygamy for both sexes to permit maximum flexibility in assignment. Allowing polygamy in COLONY permits the program to consider that a candidate parent had multiple mates when constructing pedigrees but does not force the program to assign multiple mates per parent for the final parentage assignments. Although our simulated genotypes were constructed, and so did not have any mistyping error, we set the mistyping rate to 5% to incorporate a source of uncertainty common in actual field-obtained genotypes (Table 3). We also explored the influence of mistyping rate by running a subset of analyses with a 0.5% mistyping rate, and found no meaningful differences in results (data available upon request). We set COLONY to the longest processing time permitted by the program and used the full likelihood approach to run four iterations of each analysis type to reduce sampling bias and maximize accuracy (Wang, 2016).

2.6 | Interpretation of assignments and errors

For each analysis in each program, we sorted offspring assignments into four categories: (i) correctly assigned to the true mother, (ii) incorrectly assigned to a different female, (iii) correctly left unassigned (because its mother was not included in the sample), or (iv) incorrectly left unassigned (when the true mother was included in the sample but the program failed to assign an offspring to its mother at high confidence). Each type of assignment error did not always occur but depended on the sampling context and the tactic of the true mother of the offspring, so we report results in terms of the female's reproductive tactic (nesting or parasitic). For females with relatives

in the population, errors could be biased towards relatives, so we made the distinction between incorrect assignment to relatives and nonrelatives in these cases. Error rates are presented as the proportion of total offspring that were incorrectly assigned. Comparing the results of the two simulated populations revealed few meaningful differences between the populations in the types or frequencies of errors. For simplicity, we present the average error rates for the two populations (or population-specific results are provided in Figure S1). To compare assignment errors based on confidence level, we examined both the 80% and the 95% confidence levels as cutoffs for assignment.

The proportion of offspring that can be correctly assigned to specific females depends on the sampling regime. In the analyses where all mothers are included (all candidate parents and all female), the best-case scenario would assign 100% of the offspring ($N = 480$ per population for the no-relatives analyses and $N = 720$ per population for the relatives-included analyses) to their true mother at high confidence. For the analyses where we included only nesting parents or nesting females (i.e., the parasitic mothers are missing from the sample), the best-case scenario would be to correctly assign the offspring of nesting parents to their true mother (67% of the total offspring), and correctly identify that the parasitic females were absent from the candidate parent pool and so leave their offspring unassigned (33% of the total offspring).

3 | RESULTS

3.1 | Sampling scenario

Under the best-case conditions—with all 18 loci and no relatives among the maternal candidates—we were able to identify the true mother of most of the offspring regardless of the sampling scenario (Figure 2). Errors in assignment were low (<5%) when all parental genotypes were included (All Parents), when paternal genotypes were excluded (All Females), and when genotypes of parasitic females and their mates were excluded (Nesting Parents). However, errors in assignment rose sharply (up to 19%) with the least complete sampling of the parental pool, when all paternal and parasitic female genotypes were excluded (Nesting Females). Importantly, the type of errors differed by program: CERVUS largely misassigned parasitic offspring to nesting mothers (3%–19% false assignment in CERVUS, vs. 0% in COLONY), while COLONY left some offspring of nesting mothers unassigned (2%–5% false exclusion in COLONY vs. 0% in CERVUS).

3.2 | Relatedness

When full-sib and half-sib relatives were included as maternal candidates, errors in assignment increased slightly (Figure 3). Specifically, more errors occurred (4%–5% of offspring) in the Nesting Parents scenario when using CERVUS, misassigning more parasitic offspring and leaving more nesting offspring unassigned. Also, more errors occurred (5%–10% of offspring) in the All Parents and All Females scenarios when using COLONY, leaving more nesting and parasitic offspring incorrectly unassigned. The percentage of errors in the Nesting Females scenarios remained relatively unchanged for both programs. Across the three categories of relatedness (full-sib, half-sib, or no relatives), mothers with full-sibs in the candidate pool were more likely to have misassigned (CERVUS) or unassigned (COLONY) offspring (Figure S3a,b). Notably, in the Nesting Females scenario, CERVUS misassigned 2%–19% of offspring of mothers with a full-sibling in the population ($N = 160$) to the full-sib, and 1%–14% to a nonrelative.

3.3 | Programs and confidence level

The confidence level (CL) thresholds played an important role in producing the specific types of errors we observed, especially in the two scenarios that lacked samples from parasites (Nesting Parents and Nesting Females; Figures 2 and 3). COLONY was able to accurately identify the correct mother of most of the offspring under any sampling context at the 80% CL. When we used the 95% CL as a threshold, a few assignments that COLONY made correctly at the 80% CL were now categorized (incorrectly) as unassigned offspring. CERVUS assigned more offspring correctly at the 95% CL than at the 80% CL; at the 80% CL, offspring that had been unassigned at the 95% CL threshold were now assigned to the wrong mother.

Although COLONY often lacked the power to assign an offspring at high confidence (resulting in unassigned offspring if the 95% CL is used), it nonetheless almost always indicated the correct mother when it did make assignments (Table S3). Even without parasitic female genotypes, COLONY correctly identified parasitic offspring that belonged to the same missing mother as siblings, thus inferring that they were the progeny of the same missing mother. In a small number of cases (3% or less, Figure S4a,d), COLONY made assignment errors when the parasite had a nesting full-sibling in the candidate parent pool. In contrast, CERVUS did not infer missing parental genotypes and appeared to struggle more with parasitic offspring when we excluded their mothers from the candidate parent pool, sometimes resulting in high rates of incorrectly assigned offspring (up to 33% for females with full-siblings, and 24% in females with half-siblings, Figure S3b,d).

3.4 | Number of loci

When the number of loci was reduced to nine—the average used by previous studies of CBP (Table 3) —error rates increased substantially for both programs. Error rates were highest in the All Females (5%–46% of offspring) and Nesting Females (5%–26% of offspring) scenarios (Figures 4 and 5). This was mostly due to an increase in incorrectly unassigned offspring for both programs, and an increase in misassigned offspring in CERVUS. Error rates were somewhat higher when full- and half-sib relatives were included in the maternal candidate pool (Figure 5), but error types were similar. As in the full set of loci runs, the error rate dropped with the inclusion of paternal genotypes in CERVUS only.

TABLE 3 Summary of sample regime and basic results for microsatellite-based studies of CBP in birds as reported by the authors of the studies

Species	Study	System	Loci	Number of allele (range)	Hetero-zygosity H_O	Mistyping rate	Program
Barnacle goose	Anderholm et al., 2009	CBP	10	8.79 (2–22)	0.60	3%	CERVUS
Barrow's goldeneye	Jaatinen et al., 2011	CBP	19	6.42 (2–14)	0.67	1%	CERVUS
Brant	Lemons & Sedinger, 2011	CBP	7	7.71 (2–22)	0.41	0%	CERVUS
Common eider	Tiedemann et al., 2011	CBP	7	21.86 (3–60)		3%	None
Common eider	Hario et al., 2012	CBP	10	11.50 (2–33)	0.55	NR	COLONY
Common eider	Hervey et al., 2019	CBP	11	18.00 (8–47)	0.52	15.7%	None
Mallard	Kreisinger et al., 2010	CBP, EPP	8	20.25 (4–40)	0.73	0%	CERVUS
Mandarin duck	Gong et al., 2016	CBP	8	12.88 (5–31)	0.75	1%	CERVUS
Common pochard	Šťovíček et al., 2013	CBP	17	6.06 (2–17)	0.52	NR	None
Ruddy duck	Reichart et al., 2010	CBP	10	7.60 (3–19)	0.48	0%	CERVUS
Wood duck	Nielsen et al., 2006	CBP	5	13.20 (6–25)	0.80	NR	None
Barn swallow	Petrželková et al., 2015	CBP, EPP, QP	6	20.83 (12–47)	0.82	0.6%–1%	CERVUS, COLONY
Black-capped chickadee	Otter et al., 2011	CBP, EPP, QP	3	15.67 (6–23)	0.88	1%	CERVUS
Black-headed gull	Ležalová-Piálková, 2011	CBP, EPP	6	11.33 (7–21)	0.78	NR	None
Blue tit	Griffith et al., 2009	CBP	5	8.60 (6–14)	0.75	NR	CERVUS
Burrowing owl	Rodriguez-Martínez et al., 2014	CBP, EPP	17	8.35 (3–20)	0.62	1%	CERVUS
European roller	Sánchez-Tójar et al., 2015	CBP, EPP, QP	6	4.70 (2–9)	0.35	1%	CERVUS
Western/island scrub-jay	Peer et al., 2007	CBP	7	25.0 (13–40) / 14.50 (3–17)	0.93 / 0.65	NR	CERVUS
Houbara bustard	Lesobre et al., 2010	CBP	12	7.3	0.56	1.5%	CERVUS
Imperial shag	Calderón et al., 2012	CBP, EPP	4	11.50 (7–15)	0.76	NR	None
Magellanic penguin	Marasco et al., 2020	CBP, EPP	9	10.33 (7–13)	0.69	2%	CERVUS
Monk parakeet	Martínez et al., 2013	CBP, EPP	7	8.35 (3–20)	0.62	NR	COLONY
Northern bobwhite	Davis et al., 2017	CBP, IBP	12	13.60 (5–24)	0.76	3%	COLONY
Scaled quail	Davis et al., 2017	CBP	12	8.42 (3–12)	0.79	3%	COLONY
Northern flicker	Wiebe & Kempenaers, 2009	CBP, EPP	12	16.23 (9–36)	0.75	0.03%	CERVUS
Prothonary warbler	Tucker et al., 2016	CBP	6	18.24 (10–26)	0.63	10%	CERVUS
Sage grouse	Bird et al., 2013	CBP, MP	13	4.7 (5–23)	0.68	1%	CERVUS
Eurasian hoopoe	Berthier et al., 2012	CBP, EPP	6	13.83 (9–18)	0.81	3%	CERVUS
Zebra finch	Schielzeth & Bolund, 2010	CBP	10	10.70 (5–14)	0.80	NR	None

NR = value not reported. NA = value not applicable because of the nature of the study. For the system column, we note which reproductive strategies were specifically assessed; CBP, EPP (extra-pair paternity), IBP (facultative interspecific parasitism) and QP (quasi-parasitism where the putative mother is excluded but not the putative father). For the Offspring unassigned parasitic column, C indicates that offspring were considered parasitic if unassigned but only under specific conditions, usually after applying another set of parental exclusion criteria.

[a]The percentage of parasitic young as reported by the original study. Note that some authors report the number of parasitic offspring as a percentage of assigned young only (excluding unassigned young), whereas other studies report this as a percent of *all* young genotyped, including assigned and unassigned offspring. In some cases, unassigned young were included as parasitic; in other cases, additional arguments were presented to include or exclude these offspring as parasitic or not.

4 | DISCUSSION

4.1 | Sampling context

Our simulations revealed that it is possible to accurately assign parentage even under the potentially challenging circumstances routinely observed in many species that practise conspecific brood parasitism: lack of male genotypes, unsampled parasitic females and the presence of female relatives in populations (Table 4). Nonetheless, there are some situations that could lead to substantial rates of assignment error, depending on the program and confidence threshold used. The presence of female genotypes in the

Percentage females sampled	Percentage offspring assigned	Percent offspring parasitic[a]	Nesting parasites	Non-nesting parasites	Unassigned offspring parasitic	Minimum confidence level
12	NR	12	Y	Y	Y	NA
60	80	13	Y	Y	*N*	95
NR	100	6	Y	NR	Y	NR
NA	NR	17	Y	Y	Y	NA
38–65	100	34	NR	NR	Y	NR
NR	NR	23	Y	Y	Y	NA
70	100	10.1	NA	NA	Y	95
80	100	40.9	Y	Y	C	95
100	80	39	NA	NA	Y	NR
NR	89	29	Y	Y	C	NR
NA	NR	27	NR	NR	C	NA
80	91.7	5.7	Y	NR	Y	95
85	100	55	NA	NA	*N*	80
NA	NR	9	NR	NR	*N*	95
100	89	0	NA	NA	NA	NR
100	100	5.7	NR	NR	NA	80
75	54	0	NA	NA	Y	95
NR	0	NR	*N*	*N*	NR	NR
80	73	26	*N*	Y	Y	NR
NA	NA	0	NA	NA	NA	NA
2	NR	6	NR	NR	Y	NR
100	NR	1.2	NR	NR	Y	NR
95	100	21	NR	NR	*N*	80
95	100	21	NR	NR	*N*	80
95	100	5	Y	NR	Y	95
92.5	NR	12.5	Y	Y	Y	95
20–90	47.6	2.2	NR	NR	Y	80
60	71	7	NR	NR	C	80
100	NR	5.4	Y	NA	NR	NR

candidate sample was particularly important for accurate maternity assignment. Nesting females were included in all our analyses and both programs almost always correctly assigned the offspring of these females, when the offspring were assigned (in some scenarios, offspring of nesting females that could have been assigned were left unassigned).

Offspring of females that were not included in the candidate pool were at the greatest risk of being incorrectly assigned to

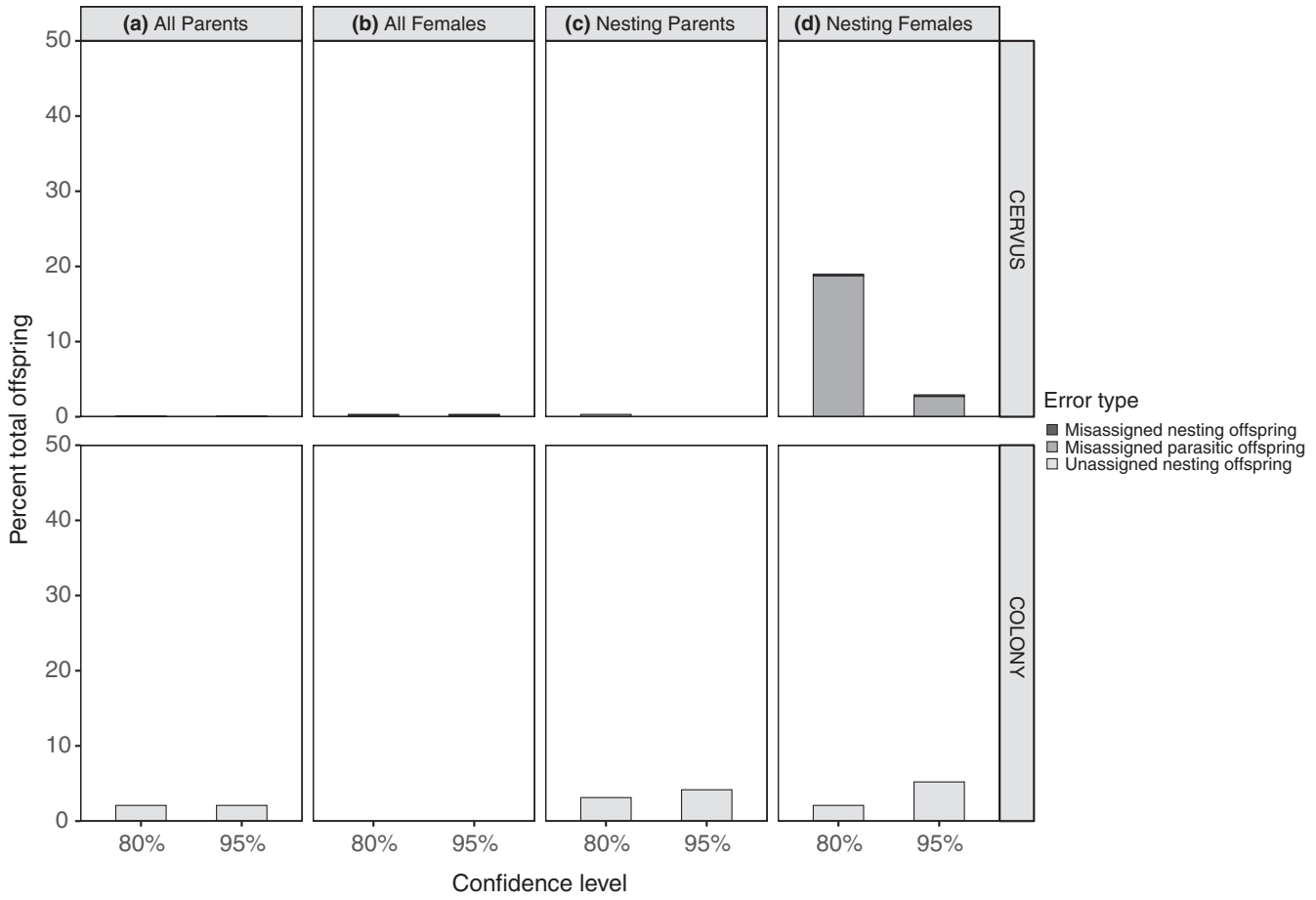

FIGURE 2 Comparison of the effect of four sampling contexts on the proportion of offspring assigned incorrectly or left unassigned by the programs COLONY and CERVUS when relatives are not included in the sample. Errors are shown for 80% or 95% confidence level cutoffs. These analyses included all 18 loci. (a) All candidate parents were included in the analysis, including parasitic females and their mates. (b) All male genotypes were excluded, so that only females were in the candidate parent pool. (c) Parasitic females and their mates were excluded from the parent pool, so that only nesting females and their mates were candidate parents. (d) Parasitic females and all males were excluded from the candidate parent pool, so that only nesting females were included. Indicated for each analysis is the percentage of the 480 total offspring (percentage averaged for the two population runs; see Table 2) that were incorrectly assigned in each of four categories of mother nesting status and error type: (i) misassigned offspring of nesting females, (ii) misassigned offspring of parasitic females, (iii) unassigned offspring of nesting females and (iv) unassigned offspring of parasitic females

another female. In some brood parasitic species, female parasites may be either nesting (laying parasitic eggs in addition to incubating their own clutch) or non-nesting (laying parasitic eggs only). In the field, genetic samples are often not obtained from non-nesting parasites, as they are less likely to be captured on a nest during a brief egg-laying visit than are incubating females. As a result, non-nesting parasites are most likely to go undetected or under-represented using genetic sampling methods alone. In the CERVUS simulations that excluded parasite genotypes, up to 19% of the total offspring were misassigned to another (nesting) female. This rate of incorrect assignment could falsely inflate rates of nesting parasitism—by overestimating parasitism from nesting females—relative to non-nesting parasitism, as well as confound estimates of individual reproductive success. Fortunately, this type of error was minimized by using the higher (95%) confidence threshold, including paternal genotypes in CERVUS, or by using COLONY assignments.

4.2 | Relatedness

The presence of relatives in the breeding population modestly increased the frequency of incorrectly unassigned offspring, indicating that high levels of kinship in a population can result in a loss of assignment power. This increase in error rates was most pronounced for CERVUS analyses using fewer loci and a strict confidence (95%) threshold, and in cases when male and/or parasitic females were not sampled. In most other cases, particularly with larger numbers of loci, both programs do well with complete sampling of candidate mother genotypes even in the face of kin structure, making very few errors. We simulated populations in which nearly half of females (n = 32 of a total 72) had a full- or half-sibling among the candidate maternal genotypes; this rate is at the high end of kin structure in our study populations (our unpublished data) and may be higher than other populations or species.

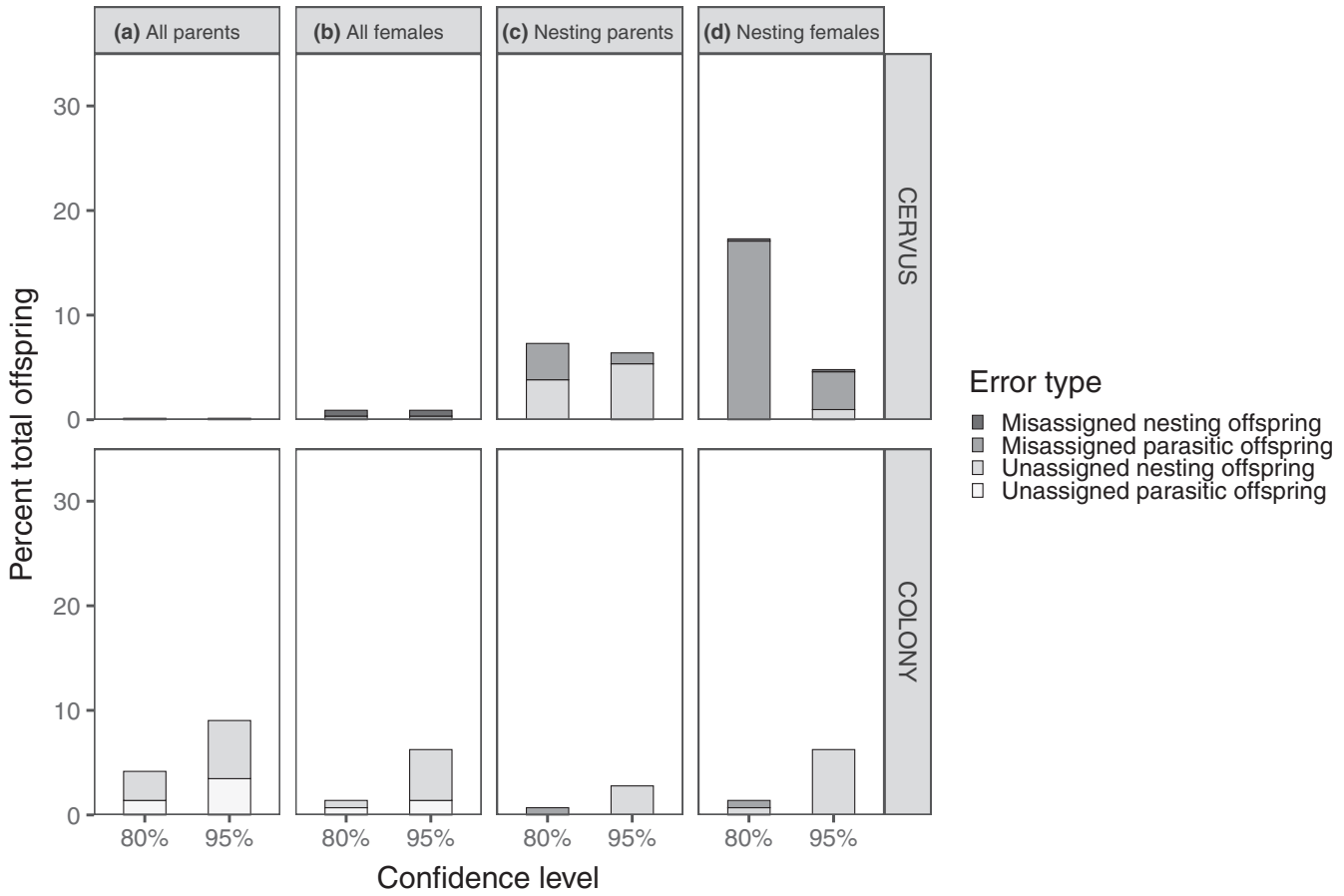

FIGURE 3 Comparison of the effect of four sampling contexts on the proportion of offspring assigned incorrectly or left unassigned by the programs COLONY and CERVUS when relatives are included in the sample. Errors are shown for 80% or 95% confidence level cutoffs. Errors are averaged for the two population simulations. These analyses included all 18 loci. (a) All candidate parents included, including parasitic females and their mates. (b) All male genotypes excluded; only females in the candidate parent pool. (c) Parasitic females and their mates excluded from the parent pool; only nesting females and their mates were candidate parents. (d) Parasitic females and all males excluded from the candidate parent pool; only nesting females included. Indicated for each analysis is the percentage of the 720 total offspring (percentage averaged for the two population runs; see Table 2) in each of four categories: (i) misassigned offspring of nesting females, (ii) mis-assigned offspring of parasitic females, (iii) unassigned offspring of nesting females, and (iv) unassigned offspring of parasitic females

A concern for CBP researchers is that kin structure could generate incorrect assignment of an offspring of a nesting female to a relative of its true mother, creating spurious cases of kin-directed parasitism. Fortunately, we found this error to be rare. Instead, our results suggest a different risk from the presence of relatives in the population: when kin were present in the candidate mother pool, more offspring were left incorrectly unassigned due to low power in discriminating among similar maternal genotypes. If unassigned offspring are considered parasitic, the presence of kin in the population can cause an inflated estimate of parasitism rate, potentially leading to false conclusions about kinship within populations and the evolution and dynamics of CBP by correlating high levels of kinship in populations to elevated parasitism rates. These problems can be avoided by not assuming that unassigned offspring are parasitic.

4.3 | Confidence levels and programs

Our simulations show that COLONY and CERVUS were prone to different types of error. Analyses using COLONY were more likely to leave offspring unassigned even when their mother was in the candidate pool (false exclusion). If unassigned offspring are considered parasitic, this type of error potentially inflates estimates of CBP. CERVUS was more likely to incorrectly assign offspring of parasites to nesting females, particularly under a relaxed (80%) confidence level, and in doing so, artificially reduces non-nesting patterns of CBP while potentially inflating rates of parasitism attributed to nesting females. Even with an imposed 80% threshold, we believe COLONY provides superior results for studies of CBP with incomplete candidate parent sampling, as the tendency to leave offspring unassigned rather than falsely assigned provides a more conservative estimate of CBP (if

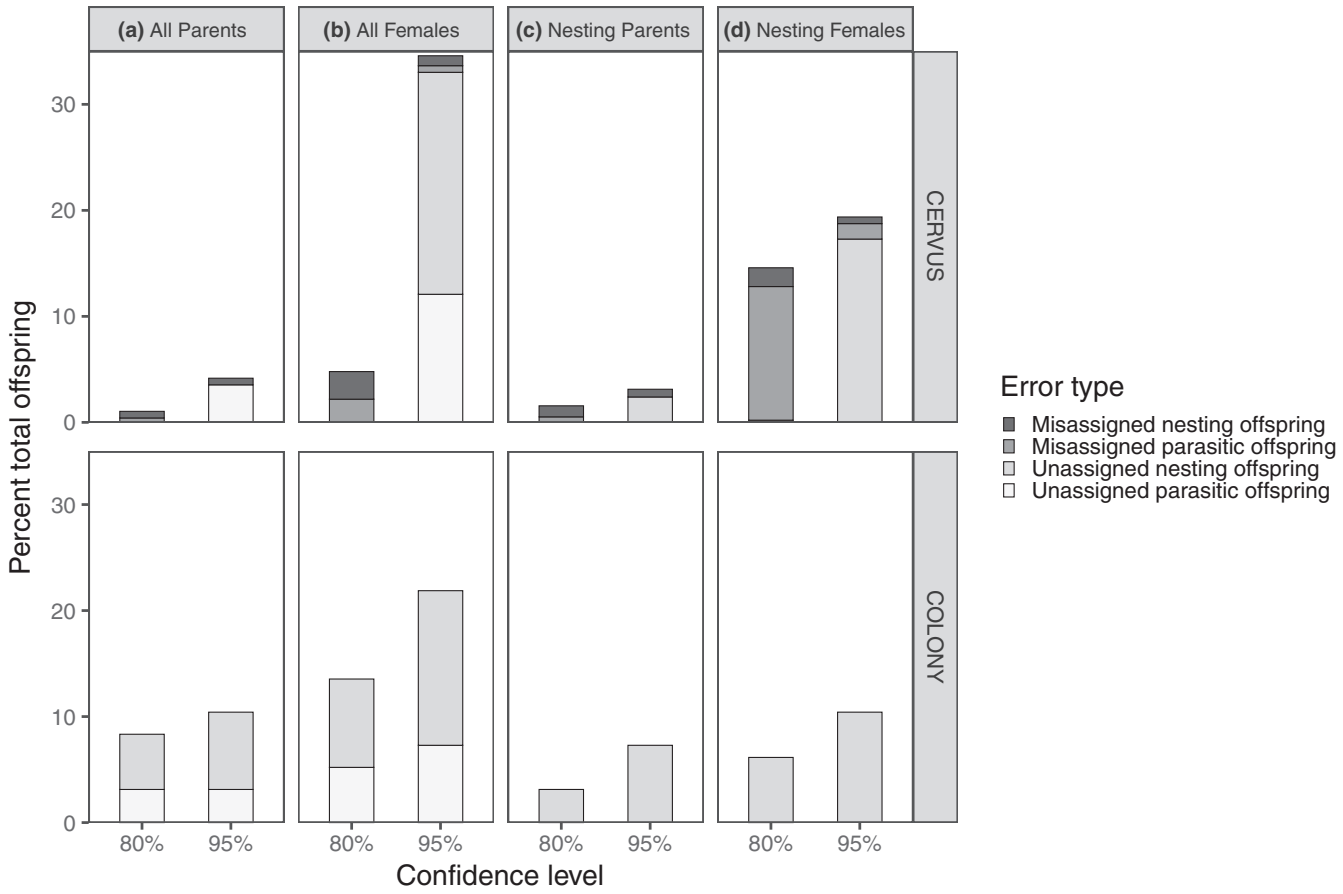

FIGURE 4 The effect of a reduced number of loci ($N = 9$) on assignment errors for different sampling scenarios varying in which parents are included in the sample of candidate parents when relatives are not included in the sample. Errors are shown for 80% or 95% confidence level cutoffs. (a) All candidate parents were included in the analysis, including parasitic females and their mates. (b) All male genotypes were excluded, so that only females were in the candidate parent pool. (c) Parasitic females and their mates were excluded from the parent pool, so that only nesting females and their mates were candidate parents. (d) Parasitic females and all males were excluded from the candidate parent pool, so that only nesting females were included. Indicated for each analysis is the percentage of the 480 total offspring (percentage averaged for the two population runs; see Table 2) in each of four categories based on nesting status of the mother and error type: (i) misassigned offspring of nesting females, (ii) misassigned offspring of parasitic females, (iii) unassigned offspring of nesting females and (iv) unassigned offspring of parasitic females

unassigned offspring are not assumed to be parasitic) and minimizes incorrect patterns of CBP. However, the best choice of program is dependent on the nature of the data available; in studies where sampling of the candidate parent pool is complete, our results suggest CERVUS could outperform COLONY, as the error rates are lower for CERVUS in that context.

A unique strength of COLONY is its ability to assign offspring to an inferred female that is missing from the population, which provides a picture of the reproductive effort and patterns of host choice of an unsampled female. While we considered the offspring that COLONY identifies as being the progeny of an unsampled mother as unassigned for the purposes of this study, a closer examination revealed that when we removed the parasitic females from the candidate parent pool, COLONY could correctly identify how many females were missing and assign all the progeny of a missing female to a unique inferred missing female identity. In other words, COLONY could identify a missing female's offspring and assign them back to her, without her in the candidate sample; this makes COLONY especially useful in studies of CBP where many of the parasites do not have nests of their own in a given year, and thus are likely to not be captured or sampled.

The two programs differed in the way that errors were attributed to one or more families. When COLONY made an error, it tended to incorrectly assign all the offspring of a particular female as a unit. In contrast, when CERVUS made errors, it tended to incorrectly assign some, but not all, of a female's offspring. Consequently, the results from COLONY are more likely to change the estimated reproductive output of a few females, whereas the results from CERVUS are more likely to change the estimated reproductive output of many females.

To be able to directly compare results between COLONY and CERVUS, we imposed confidence level thresholds for COLONY based on CERVUS standards (80% and 95%). However, these two programs calculate and report confidence in assignment in different ways, and the imposed confidence levels we used sometimes changed how we interpreted assignments in COLONY. For example, the few (up to 5%, depending on the analysis) COLONY assignments made at confidence

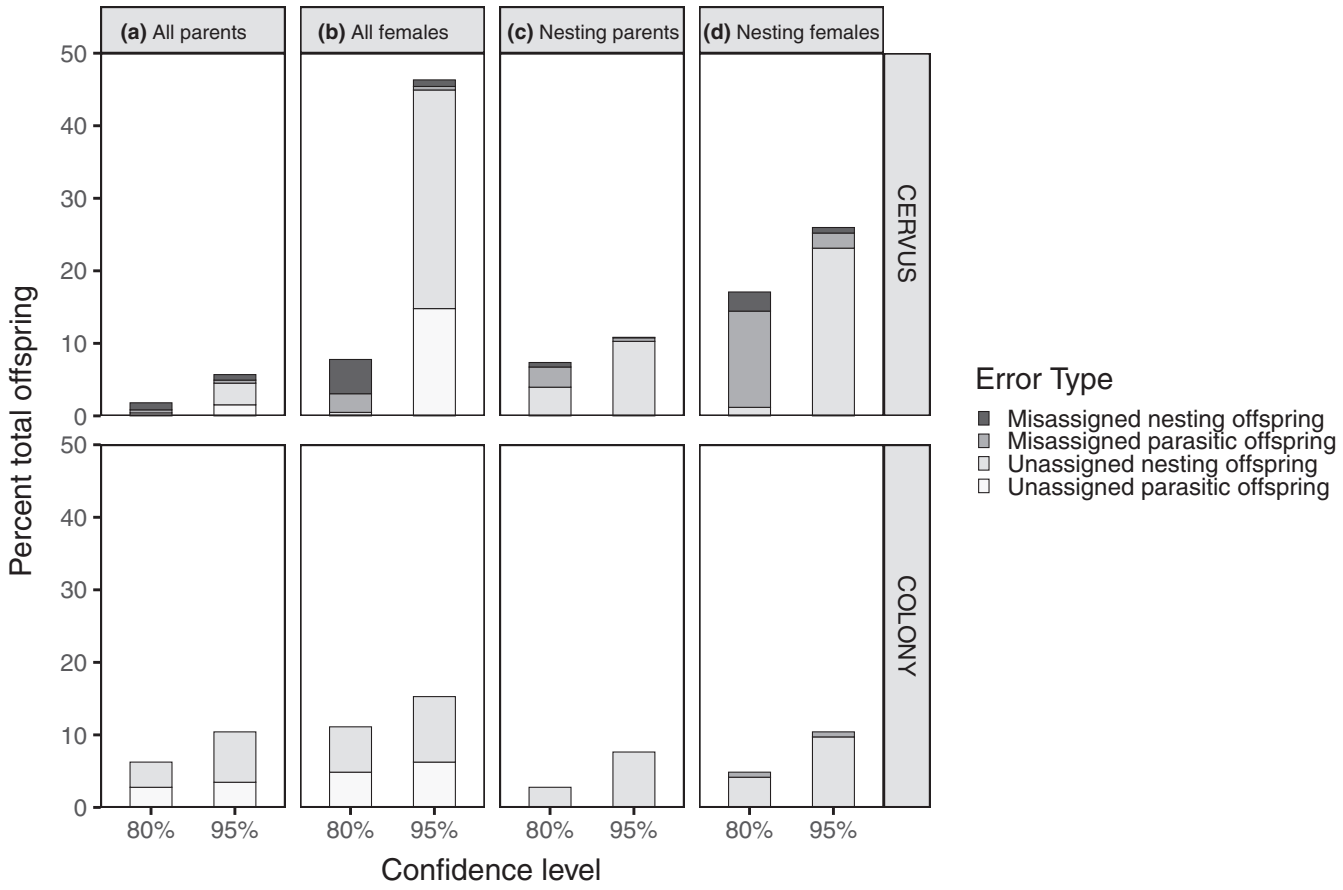

FIGURE 5 The effect of a reduced number of loci ($N = 9$) on assignment errors for different sampling scenarios varying in which parents are included in the sample of candidate parents when relatives are included in the sample. Errors are shown for both an 80% and 95% confidence level cutoff. Errors are averaged for the two population simulations. (a) All candidate parents included including parasitic females and their mates. (b) All male genotypes excluded; only females included. (c) Parasitic females and their mates excluded; only nesting females included. (d) Parasitic females and males excluded; only nesting females included. Within each panel, errors in assignment are sorted by the reproductive tactic of the mother, so offspring are either (i) misassigned offspring of nesting females, (ii) misassigned offspring of parasitic females, (iii) unassigned offspring of nesting females and (iv) unassigned offspring of parasitic females

TABLE 4 Summary of the programs COLONY and CERVUS and their performance given the challenges presented in this study

Program details	CERVUS	COLONY
Year released	1998	2010
Assignment approach	Pairwise maximum likelihood	Pedigree maximum likelihood
Confidence level indicator	LOD score	Probability of relationship in assessed pedigree options
Performance		
Unsampled candidate parents	High risk of incorrect assignment when few loci included (particularly with low allelic diversity/polymorphic information content)	Low risk of incorrect assignment when parent is present in the candidate parent pool. COLONY can also infer and assign to missing parents with high accuracy.
Related candidate parents	Risk of incorrect assignment, particularly with few loci and 95% confidence threshold	Risk of not assigning offspring to any parent, particularly with few loci
Small number of loci	Risk of incorrect assignment	Risk of not assigning offspring to any parent
Confidence level	95% confidence level improves accuracy with more loci and/or complete parent pool but risks leaving offspring incorrectly unassigned, 80% confidence level risks incorrect assignment	Calculated differently than CERVUS, and is not correlated strongly with accuracy, so it is safe to accept assignments made at any probability.

levels lower than 80% were generally correct (Table S3) but with the imposed 80% confidence level threshold these assignments were considered incorrectly unassigned. Few studies that use COLONY either report or impose a confidence level threshold (Table 3), and the fact that we did so to match CERVUS's assignment system meant that our interpretation of the apparent accuracy of COLONY assignments was probably conservative. Accordingly, we suggest not imposing a confidence level threshold on COLONY assignments to eliminate incorrect lack of assignment.

4.4 | Number of loci

Not surprisingly, reducing the number of loci in our simulations dramatically increased the number of offspring that were incorrectly left unassigned, even when all maternal genotypes were included in the candidate pool. Including male genotypes in the candidate parent pool reduced both incorrect assignments and incorrect lack of assignments, apparently compensating in part for the lack of assignment power due to the reduced number of loci included in these analyses.

The implications of incorrectly leaving offspring unassigned depends on the assumptions that researchers make. Some previous studies assumed that lack of assignment indicates that an offspring is the product of a parasitic event or extra-pair mating (e.g., Lemons & Sedinger, 2011; Lesobre et al., 2010; Tucker et al., 2016). Studies with sampling contexts that are prone to this error—those with few markers and a lack of samples from sires—should avoid assuming that unassigned offspring are parasitic.

4.5 | Implications for past and future field studies

To put our analyses into broader context, we surveyed 28 studies that used genetic parentage assignment methods to draw inferences about conspecific brood parasitism. This sample reveals considerable variation, both in the approaches taken and in the conditions that could affect error rates (Table 3). As CERVUS was developed 12 years before COLONY, it is not surprising that the majority of the studies used the program CERVUS (17 of 28) to assign parentage, while three used COLONY, and one study used both. Several studies (seven) did not use either parentage assignment program but instead relied on methods such as comparing the genotypes of females and offspring and assigning maternity based on the occurrence of mismatches (often only at one or two loci) relative to that expected given allele frequencies, expected mutation rates, genotyping error, null alleles or allele drop-out (e.g., Hervey et al., 2019; Šťovíček et al., 2013; Tiedemann et al., 2011). The choice to use exclusion by mismatch instead of parentage assignment may reflect a discomfort with available programs, including the limitations that we have tried to address in the present study (Anderholm et al., 2009); however, despite the appeal of simplicity, exclusion is somewhat arbitrary (Flanagan & Jones, 2019). The wide range of approaches used suggests some re-evaluation of the efficacy and accuracy of these approaches is warranted.

Our results, specifically, suggest that we need to reconsider some of the results based on analyses using CERVUS to draw conclusions about parentage and subsequent fitness proxies. Several studies used an 80% confidence level cutoff in CERVUS combined with a small number of loci, a combination that comes with a substantial risk of error (Figures 4 and 5). Indeed, of the 17 studies employing CERVUS, 10 used nine or fewer loci (in total, 15 of the 28 studies used nine or fewer loci; Table 3). Understandably, the number of microsatellite markers available has increased over the past decade, but caution is clearly warranted for studies using few loci or loci with low allelic diversity or reliability. Other studies employed a more extensive set of loci (>15 loci) for assignment, which reduces the risk of incorrect lack of assignment regardless of the program used (Jaatinen et al., 2011; Rodriguez-Martínez et al., 2014). However, even with a 95% confidence level cutoff, the error rates we observed in our analyses using CERVUS could influence the results and interpretation of some studies (Indykiewicz et al., 2017; Lemons & Sedinger, 2011). In some cases, reported levels of parasitism for some species could simply be the result of assignment error and falsely indicate the presence of CBP in a species where it does not actually occur.

These concerns are further compounded by the fact that there is extensive variation among studies of CBP in birds (Table 3) in the number of alleles, observed heterozygosity (H_O), excess or deficiency in heterozygosity ($H_E - H_O$), and polymorphic information content (PIC) for the microsatellite markers used (Figure S4; Table S4). Markers with few alleles or with very uneven distributions of alleles can strongly influence the power of the parentage analyses. Although varying the number of alleles or heterozygosity was beyond the scope of our study, had we conducted simulations using loci with fewer alleles or lower degrees of variability, the results would have been even more dramatic, indicating that parentage assignments under such circumstance are unreliable. In our analyses, the average number of alleles was 10.5 (range 3–36), and average observed heterozygosity was 0.67 (range 0.21–0.96; Figure S4a,b), directly in line with the studies summarized in Table 3 (average number of alleles 11.7 [range 2–60] and average observed heterozygosity 0.69 [range 0.00–1.00]; Figure S4). Hence, the results of our simulations, including those with a reduced number of loci, are representative and applicable to other studies of CBP.

Our results are also relevant to other social systems beyond CBP. The ability to correctly identify biological parents is essential in other social and mating systems in which broods of mixed parentage arise. Extra-pair paternity (EPP) is common in birds and is the male analogue to analyses of CBP in that exclusion of the social male as the parent of a nest implicates the occurrence of EPP, whereby a female copulates outside the pair bond so that some of the offspring of a pair are sired by a male other than the social partner (Westneat & Stewart, 2003). As with CBP, false exclusion of the social male when he is actually the sire of an offspring in his nest leads to the incorrect conclusion that an interesting biological phenomenon has occurred—EPP. Admittedly, the constellation of factors that contribute to errors

in assessing CBP, such as relatedness and lack of samples from the parent of primary interest (in this case, males), is less of an issue in studies of EPP because of a focus on capturing males and limited or no relatedness typically among neighbouring males (but see Mulder, 1995 and Colombelli-Négrel et al., 2009 for a possible example in male superb fairy-wrens). Moreover, our study shows that even with a reduced number of loci, the risk of false exclusions should be low under the usual sampling circumstances (i.e., samples from most males, low relatedness). However, one mating system where false exclusion is likely to be important is quasi-parasitism. This system combines CBP with EPP, whereby a male copulates with a conspecific brood parasite and sires the "parasitic" eggs she then lays in his nest. The genetic signature for quasiparasitism is exclusion of the female but not the male as parents of one or more offspring in a nest. This is exactly the pattern generated by false exclusion of a true social mother (Griffith et al., 2004). Quasiparasitism should be very rare in nature because it is unclear why males would benefit from the quasiparasitism and the probable costs should exceed the scant benefits (Lyon et al., 2002). Given this, many if not most of the rare putative examples of quasiparasitism in the literature may be the result of false exclusions of the true mother of offspring (Griffith et al., 2004).

Our results highlight the consequences of incomplete sampling of parents and the choice of confidence levels on accurate assignment; yet, these parameters have not always been reported in the literature on CBP. Our study specifically examined the consequences of missing male genotypes or a large portion of candidate females (e.g., females that are parasitic only and hence unlikely to be caught on the nest). Often researchers typically assume some proportion of females are included in the candidate pool; of the 17 studies using CERVUS (Table 3), six sampled less than 75% of the female candidate pool, and another three did not report the percentage sampled. One of three studies using COLONY sampled less than 75% of the female candidate pool. Regarding confidence level, only six of the 17 studies employing CERVUS used a confidence level of 95%, four used a confidence level of 80% and seven other studies did not report any confidence level. Two of three studies using COLONY did not report a confidence level or probability of assignment, although as noted, COLONY does not use these probabilities as thresholds for assignment in the same way as CERVUS. Of all 28 studies, only 13 reported any confidence level.

The use of microsatellites for parentage assignment may continue to decline as SNP-based approaches become more accessible and affordable (Flanagan & Jones, 2019). Yet, small numbers of microsatellites have comparable statistical power to SNP panels (reviewed in Flanagan & Jones, 2019), and are likely to be used for compatibility with historical data sets. Additionally, parentage assignment using SNP genotypes is also conducted with CERVUS (e.g., Cramer et al., 2011; Kaiser et al., 2017) and COLONY (e.g., Weinman et al., 2015), and similar methodological concerns about parent sampling and kin structure apply (Flanagan & Jones, 2019).

Our point is not to find fault with previous studies nor with the utility of any program specifically. Rather, we simply emphasize that choice of program, number of loci and the composition of the candidate pool can have considerable impact on the accuracy of parentage assignments. This is not new or surprising, but the variation observed among studies of CBP does suggest that we need to evaluate these results carefully, and researchers need to take greater care to ensure that their assignments are robust by evaluating their experimental design, field data collection and subsequent analyses in the context of the strengths and weaknesses of genetic approaches. This is particularly true when there are multiple constraints on the accuracy of parentage assignments, such as an inability to sample males, incomplete sampling of the female candidate pool, and the possibility of relatives in the population. Having undertaken such an assessment for our population of wood ducks, we are reassured that it is possible to assign maternity with confidence even under these circumstances, provided the number of loci is reasonable (>15) and the loci are sufficiently variable, and by employing COLONY, whose pedigree approach to assign parentage and ability to infer unsampled parents makes it particularly suitable for studies of alternative reproductive tactics. We caution that these results may be specific to our population and the microsatellite loci we used. Accordingly, we urge researchers to undertake comparable evaluations when conducting studies under similar—but not uncommon—circumstances. Without more rigorous assessments, we risk creating enticing adaptive explanations of behaviours that may instead simply represent artefacts of inaccurate parentage assignments.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers for their comments and suggestions that greatly improved the manuscript, as well as Tez Stair, Melissa Jones, Mitch Hinton and many field assistants and student interns for long hours in the field collecting data and samples. We also thank Maria Cecilia Penedo and her staff at the Veterinary Genetics Lab (VGL) at UC Davis for microsatellite genotyping of our wood duck populations. Scott Blackwell contributed to the illustrations used in Figure 1. This study was approved by the Animal Care and Use Committee at UC Davis (Protocol # 7535 and 20971). Funding was provided by the National Science Foundation NSF IOS 1355208 (J.M.E.) and NSF IOS 1354894 (B.E.L.) and the D. G. Raveling Waterfowl Endowment (J.M.E.). This work was also supported by Project No. CA-D-WFB-6342-H/Accession no. 1013603 from the USDA National Institute of Food and Agriculture (J.M.E.). Additional graduate student support was provided to C.T. from UC Santa Cruz and post-doctoral support from the D. G. Raveling Waterfowl Endowment to C.W.

AUTHOR CONTRIBUTIONS

All of the authors designed the study and contributed substantially to the conceptual framework. C.W. and B.L. constructed the simulated genotypes. C.T. conducted the maternity and paternity analyses, summarized the data, and created most graphs. The literature review and synthesis were conducted by C.T., J.E. and C.W. The initial draft of the manuscript was written by C.T. with input from all authors, and all authors edited, wrote new sections, and contributed to multiple revisions of the manuscript.

DATA AVAILABILITY STATEMENT

Simulated wood duck genotypes: Dryad https://doi.org/10.7291/D1067B. Maternity analysis results from COLONY and CERVUS for the simulated wood duck populations: Dryad https://doi.org/10.7291/D1VH5F. Thow et al., 2021a,b.

ORCID

Caroline M. Thow https://orcid.org/0000-0001-6630-0282
Caitlin P. Wells https://orcid.org/0000-0002-8840-3095

REFERENCES

Åhlund, M., & Andersson, M. (2001). Female ducks can double their reproduction. *Nature*, *414*, 600–601. https://doi.org/10.1038/414600b

Andersson, M. (1984). Brood parasitism within species In C. J. Barnard (Ed.), *Producers and scroungers: strategies of exploitation and parasitism* (pp. 195–228). Croom-Helm.

Anderholm, S., Marshall, R. C., van der Jeugd, H. P., Waldeck, P., Larsson, K., & Andersson, M. (2009). Nest parasitism in the barnacle goose: evidence from protein fingerprinting and microsatellites. *Animal Behaviour*, *78*(1), 167–174. https://doi.org/10.1016/j.anbehav.2009.04.011

Andersson, M., Åhlund, M., & Waldeck, P. (2019). Brood parasitism, relatedness and sociality: A kinship role in female reproductive tactics. *Biological Reviews*, *94*(1), 307–327. https://doi.org/10.1111/brv.12455

Araki, H., & Blouin, M. S. (2005). Unbiased estimation of relative reproductive success of different groups: Evaluation and correction of bias caused by parentage assignment errors. *Molecular Ecology*, *14*(13), 4097–4109. https://doi.org/10.1111/j.1365-294X.2005.02689.x

Arnold, K. E., & Owens, I. P. F. (2002). Extra-pair paternity and egg dumping in birds: Life history, parental care and the risk of retaliation. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1497), 1263–1269. https://doi.org/10.1098/rspb.2002.2013

Avise, J. C., Jones, A. G., Walker, D., & DeWoody, J. A. (2002). Genetic mating systems and reproductive natural histories of fishes: Lessons for ecology and evolution. *Annual Review of Genetics*, *36*(1), 19–45. https://doi.org/10.1146/annurev.genet.36.030602.090831

Baldassarre, G. (2014). *Ducks, geese, and swans of North America* (Vol. *1*). Johns Hopkins University Press.

Bellrose, F., & Holm, D. (1994). *Ecology and management of the wood duck* (1st ed.). Stackpole Books.

Berger-Wolf, T. Y., Sheikh, S. I., DasGupta, B., Ashley, M. V., Caballero, I. C., Chaovalitwongse, W., & Putrevu, S. L. (2007). Reconstructing sibling relationships in wild populations. *Bioinformatics*, *23*(13), 49–56. https://doi.org/10.1093/bioinformatics/btm219

Berthier, K., Leippert, F., Fumagalli, L., & Arlettaz, R. (2012). Massive nest-box supplementation boosts fecundity, survival and even immigration without altering mating and reproductive behaviour in a rapidly recovered bird population. *PLoS One*, *7*(4), 2–7. https://doi.org/10.1371/journal.pone.0036028

Bird, K. L., Aldridge, C. L., Carpenter, J. E., Paszkowski, C. A., Boyce, M. S., & Coltman, D. W. (2013). The secret sex lives of sage-grouse: Multiple paternity and intraspecific nest parasitism revealed through genetic analysis. *Behavioural Ecology*, *24*(1), 29–38. https://doi.org/10.1093/beheco/ars132

Brouwer, L., & Griffith, S. C. (2019). Extra-pair paternity in birds. *Molecular Ecology*, *28*(22), 4864–4882. https://doi.org/10.1111/mec.15259

Buchholz, W. G., Pearce, J. M., Pierson, B. J., & Scribner, K. T. (1998). Dinucleotide repeat polymorphisms in waterfowl (family Anatidae): Characterization of a sex-lined (Z-specific) and 14 autosomal loci. *Animal Genetics*, *29*, 323–325.

Calderón, L., Svagelj, W. S., Quintana, F., Lougheed, S. C., & Tubaro, P. L. (2012). No evidence of extra-pair paternity or intraspecific brood parasitism in the Imperial Shag *Phalacrocorax atriceps*. *Journal of Ornithology*, *153*(2), 399–404. https://doi.org/10.1007/s10336-011-0754-6

Colombelli-Négrel, D., Schlotfeldt, B. E., & Kleindorfer, S. (2009). High levels of extra-pair paternity in Superb Fairy-wrens in South Australia despite low frequency of auxiliary males. *Emu – Austral Ornithology*, *109*(4), 300–304. https://doi.org/10.1071/MU09035

Cramer, E. R. A., Hall, M. L., de Kort, S. R., Lovette, I. J., & Vehrencamp, S. L. (2011). Infrequent extra-pair paternity in the banded wren, a synchronously breeding tropical passerine. *The Condor*, *113*(3), 637–645. https://doi.org/10.1525/cond.2011.100233

Davis, C. A., Orange, J. P., Van Den Bussche, R. A., Elmore, R. D., Fuhlendorf, S. D., Carroll, J. M., Tanner, E. P., & Leslie, D. M. (2017). Extrapair paternity and nest parasitism in two sympatric quail. *The Auk*, *134*(4), 811–820. https://doi.org/10.1642/AUK-16-162.1

Denk, A. G., Gautschi, B., Carter, K., & Kempenaers, B. (2004). Seven polymorphic microsatellite loci for paternity assessment in the mallard (*Anas platyrhynchos*). *Molecular Ecology Notes*, *4*(3), 506–508. https://doi.org/10.1111/j.1471-8286.2004.00707.x

Double, M. C., Cockburn, A., Barry, S. C., & Smouse, P. E. (1997). Exclusion probabilities for single-locus paternity analysis when related males compete for matings. *Molecular Ecology*, *6*(12), 1155–1166. https://doi.org/10.1046/j.1365-294X.1997.00291.x

Eadie, J. M. (1989). *Alternative reproductive tactics in a precocial bird: The ecology and evolution of brood parasitism in goldeneyes*. University of British Columbia.

Eadie, J. M., Kehoe, F. P., & Nudds, T. D. (1988). Pre-hatch and post-hatch brood amalgamation in North American Anatidae: A review of hypotheses. *Canadian Journal of Zoology*, *66*(8), 1709–1721. https://doi.org/10.1139/z88-247

Eadie, J. M., Smith, J. N. M., Zadworny, D., Kühnlein, U., & Cheng, K. (2010). Probing parentage in parasitic birds: An evaluation of methods to detect conspecific brood parasitism using goldeneyes *Bucephala islandica* and *B. clangula* as a test case. *Journal of Avian Biology*, *41*(2), 163–176. https://doi.org/10.1111/j.1600-048X.2009.04735.x

Fields, R. L., & Scribner, K. T. (1997). Isolation and characterization of novel waterfowl microsatellite loci: Cross-species comparisons and research applications. *Molecular Ecology*, *6*(2), 199–202. https://doi.org/10.1046/j.1365-294X.1997.d01-123.x

Flanagan, S. P., & Jones, A. G. (2019). The future of parentage analysis: From microsatellites to SNPs and beyond. *Molecular Ecology*, *28*, 544–567.

Forslund, P. Ä. R., & Larsson, K. (1995). Intraspecific nest parasitism in the barnacle goose: Behavioural tactics of parasite and host. *Animal Behaviour*, *50*, 509–517.

Gong, Y., Kimball, R. T., Mary, C. S., Cui, X., Wang, L., Jiang, Y., & Wang, H. (2016). Kin-biased conspecific brood parasitism in a native Mandarin duck population. *Journal of Ornithology*, *157*(4), 1063–1072. https://doi.org/10.1007/s10336-016-1353-3

Griffith, S. C., Barr, I., Sheldon, B. C., Rowe, L. V., & Burke, T. (2009). Egg patterning is not a reliable indicator of intraspecific brood parasitism in the blue tit *Cyanistes caeruleus*. *Journal of Avian Biology*, *40*(3), 337–341. https://doi.org/10.1111/j.1600-048X.2009.04671.x

Griffith, S. C., Lyon, B. E., & Montgomerie, R. (2004). Quasi-parasitism in birds. *Behavioral Ecology and Sociobiology*, *56*, 191–200. https://doi.org/10.1007/s00265-004-0766-9

Griffith, S. C., Owens, I. P., & Thuman, K. A. (2002). Extra pair paternity in birds: A review of interspecific variation and adaptive function. *Molecular Ecology*, *11*, 2195–2212. https://doi.org/10.1046/j.1365-294X.2002.01613.x

Guerier, A. S., Bishop, J. M., Crawford, S. J., Schmidt-Küntzel, A., & Stratford, K. J. (2012). Parentage analysis in a managed free ranging population of southern white rhinoceros: Genetic diversity, pedigrees and management. *Conservation Genetics*, *13*(3), 811–822. https://doi.org/10.1007/s10592-012-0331-4

Hario, M., Koljonen, M. L., & Rintala, J. (2012). Kin structure and choice of brood care in a common eider (*Somateria mollissima*) population.

Journal of Ornithology, *153*(3), 963–973. https://doi.org/10.1007/sl0336-012-0825-3

Harrison, H. B., Saenz-Agudelo, P., Planes, S., Jones, G. P., & Berumen, M. L. (2013). Relative accuracy of three common methods of parentage analysis in natural populations. *Molecular Ecology*, *22*(4), 1158–1170. https://doi.org/10.1111/mec.12138

Hervey, S. D., Barnas, A. F., Stechmann, T. J., Rockwell, R. F., Ellis-Felege, S. N., & Darby, B. J. (2019). Kin grouping is insufficient to explain the inclusive fitness gains of conspecific brood parasitism in the common eider. *Molecular Ecology*, *28*(21), 4825–4838. https://doi.org/10.1111/mec.15258

Hoffman, J. I., & Amos, W. (2005). Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, *14*(2), 599–612. https://doi.org/10.1111/j.1365-294X.2004.02419.x

Indykiewicz, P., Podlaszczuk, P., & Minias, P. (2017). Extra-pair paternity in the black-headed gull: Is it exceptional among colonial waterbirds? *Behaviour*, *154*(11), 1081–1099. https://doi.org/10.1163/1568539X-00003459

Jaatinen, K., Jaari, S., O'Hara, R. B., Ost, M., & Merilä, J. (2009). Relatedness and spatial proximity as determinants of host-parasite interactions in the brood parasitic Barrow's goldeneye (*Bucephala islandica*). *Molecular Ecology*, *18*(12), 2713–2721. https://doi.org/10.1111/j.1365-294X.2009.04223.x

Jaatinen, K., Öst, M., Gienapp, P., & Merilä, J. (2011). Differential responses to related hosts by nesting and non-nesting parasites in a brood-parasitic duck. *Molecular Ecology*, *20*(24), 5328–5336. https://doi.org/10.1111/j.1365-294X.2011.05281.x

Jones, A. G., Small, C. M., Paczolt, K. A., & Ratterman, N. L. (2010). A practical guide to methods of parentage analysis. *Molecular Ecology Resources*, *10*(1), 6–30. https://doi.org/10.1111/j.1755-0998.2009.02778.x

Jones, O. R., & Wang, J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, *10*(3), 551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Kaiser, S. A., Taylor, S. A., Chen, N., Sillett, T. S., Bondra, E. R., & Webster, M. S. (2017). A comparative assessment of SNP and microsatellite markers for assigning parentage in a socially monogamous bird. *Molecular Ecology Resources*, *17*(2), 183–193. https://doi.org/10.1111/1755-0998.12589

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*(5), 1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

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*(2), 576–579. https://doi.org/10.1111/j.1471-8286.2006.01256.x

Karaket, T., & Poompuang, S. (2012). CERVUS vs. COLONY for successful parentage and sibship determinations in freshwater prawn *Macrobrachium rosenbergii de Man*. *Aquaculture*, *324–325*, 307–311. https://doi.org/10.1016/j.aquaculture.2011.10.045

Kreisinger, J., Munclinger, P., Javůrková, V., & Albrecht, T. (2010). Analysis of extra-pair paternity and conspecific brood parasitism in mallards (*Anas platyrhynchos*) using non-invasive techniques. *Journal of Avian Biology*, *41*(5), 551–557. https://doi.org/10.1111/j.1600-048X.2010.05002.x

Lemons, P. R., Marshall, T. C., McCloskey, S. E., Sethi, S. A., Schmutz, J. A., & Sedinger, J. S. (2014). A likelihood-based approach for assessment of extra-pair paternity and conspecific brood parasitism in natural populations. *Molecular Ecology Resources*, *15*, 1–10. https://doi.org/10.1111/1755-0998.12287

Lemons, P. R., & Sedinger, J. S. (2011). Egg size matching by an intraspecific brood parasite. *Behavioural Ecology*, *22*(4), 696–700. https://doi.org/10.1093/beheco/arr035

Lesobre, L., Lacroix, F., Le Nuz, E., Hingrat, Y., Chalah, T., & Jaime, M. S. (2010). Absence of male reproductive skew, along with high frequency of polyandry and conspecific brood parasitism in the lekking Houbara bustard *Chlamydotis undulata undulata*. *Journal of Avian Biology*, *41*(2), 117–127. https://doi.org/10.1111/j.1600-048X.2009.04794.x

Ležalová-Piálková, R. (2011). Molecular evidence for extra-pair paternity and intraspecific brood parasitism in the black-headed gull. *Journal of Ornithology*, *152*(2), 291–295. https://doi.org/10.1007/s10336-010-0581-1

Lyon, B. E., & Eadie, J. M. (2008). Conspecific brood parasitism in birds: A life-history perspective. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 343–363. https://doi.org/10.1146/annurev.ecolsys.39.110707.173354

Lyon, B. E., & Eadie, J. M. (2017). Why do birds lay eggs in conspecifics' nests? In M. Soler (Ed.), *Avian brood parasitism – Behaviour, Ecology, evolution and coevolution* (pp. 1–33). Springer Nature.

Lyon, B. E., Hochachka, W. M., & Eadie, J. M. (2002). Paternity-parasitism trade-offs: A model and test of host-parasite cooperation in an avian conspecific brood parasite. *Evolution*, *56*, 1253–1266. https://doi.org/10.1111/j.0014-3820.2002.tb01436.x

Maak, S., Neumann, K., von Lengerken, G., & Gattermann, R. (2000). First seven microsatellites developed for the Peking duck (*Anas platyrhynchos*). *Animal Genetics*, *31*, 233.

Maak, S., Wimmers, K., Weigend, S., & Neumann, K. (2003). Isolation and characterization of 18 microsatellites in the Peking duck (*Anas platyrhynchos*) and their application in other waterfowl species. *Molecular Ecology Notes*, *3*(2), 224–227. https://doi.org/10.1046/j.1471-8286.2003.00405.x

Macwhirter, R. B. (1989). Minireview: On the rarity of intraspecific brood parasitism. *The Condor*, *91*(2), 485–492.

Marasco, A. C. M., Morgante, J. S., Barrionuevo, M., Frere, E., & de Mendonça Dantas, G. P. (2020). Molecular evidence of extra-pair paternity and intraspecific brood parasitism by the Magellanic penguin (*Spheniscus magellanicus*). *Journal of Ornithology*, *161*(1), 125–135. https://doi.org/10.1007/s10336-019-01720-4

Martínez, J. J., De Aranzamendi, M. C., Masello, J. F., & Bucher, E. H. (2013). Genetic evidence of extra-pair paternity and intraspecific brood parasitism in the monk parakeet. *Frontiers in Zoology*, *10*(1), 1–7. https://doi.org/10.1186/1742-9994-10-68

McRae, S. B., & Burke, T. (1996). Intraspecific brood parasitism in the moorhen: Parentage and parasite-host relationships determined by DNA fingerprinting. *Behavioral Ecology and Sociobiology*, *38*(2), 115–129. https://doi.org/10.1007/s002650050224

Mulder, R. (1995). Natal and breeding dispersal in a co-operative, extra-group-mating bird. *Journal of Avian Biology*, *26*(3), 234–240. https://doi.org/10.2307/3677324

Nielsen, C. R., Parker, P. G., & Gates, R. J. (2006). Intraspecific nest parasitism of cavity-nesting wood ducks: Costs and benefits to hosts and parasites. *Animal Behaviour*, *72*(4), 917–926. https://doi.org/10.1016/j.anbehav.2006.03.004

Nielsen, R., Mattila, D. K., Clapham, P. J., & Palsbøll, P. J. (2001). Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. *Genetics*, *157*(4), 1673–1682. https://doi.org/10.1093/genetics/157.4.1673

Odell, N. S. (2008). *Female reproductive investment in a conspecific brood parasite*. University of California.

Odell, N. S., & Eadie, J. M. (2010). Do wood ducks use the quantity of eggs in a nest as a cue to the nest's value? *Behavioural Ecology*, *21*, 794–801. https://doi.org/10.1093/beheco/arq055

Olsen, J. B., Busack, C., Britt, J., & Bentzen, P. (2001). The aunt and uncle effect: an empirical evaluation of the confounding influence of full sibs of parents on pedigree reconstruction. *The Journal of Heredity*, *92*(3), 243–247. https://doi.org/10.1093/jhered/92.3.243

Otter, K. A., Murray, B. W., Holschuh, C. I., & Fort, K. T. (2011). Rare insights into intraspecific brood parasitism and apparent

quasi-parasitism in black-capped chickadees. *Animal Biodiversity and Conservation*, *34*(1), 23–29.

Paulus, K. B., & Tiedemann, R. (2003). Ten polymorphic autosomal microsatellite loci for the eider duck (*Somateria mollissima*) and their cross-species applicability among waterfowl species (Anatidae). *Molecular Ecology Notes*, *3*(2), 250–252. https://doi.org/10.1046/j.1471-8286.2003.00414.x

Peer, B. D., Rothstein, S. I., Delaney, K. S., & Fleischer, R. C. (2007). Defence behaviour against brood parasitism is deeply rooted in mainland and island scrub-jays. *Animal Behaviour*, *73*(1), 55–63. https://doi.org/10.1016/j.anbehav.2006.06.005

Petrie, M., & Møller, A. P. (1991). Laying eggs in others' nests: intraspecific brood parasitism in birds. *Trends in Ecology and Evolution*, *6*(10), 315–320. https://doi.org/10.1016/0169-5347(91)90038-Y

Petrželková, A., Michálková, R., Albrechtová, J., Cepák, J., Honza, M., Kreisinger, J., Munclinger, P., Soudková, M., Tomášek, O., & Albrecht, T. (2015). Brood parasitism and quasi-parasitism in the European barn swallow *Hirundo rustica rustica*. *Behavioral Ecology and Sociobiology*, *69*(9), 1405–1414. https://doi.org/10.1007/s00265-015-1953-6

Reichart, L. M., Anderholm, S., Muñoz-Fuentes, V., & Webster, M. S. (2010). Molecular identification of brood-parasitic females reveals an opportunistic reproductive tactic in ruddy ducks. *Molecular Ecology*, *19*(2), 401–413. https://doi.org/10.1111/j.1365-294X.2009.04435.x

Rodriguez-Martínez, S., Carrete, M., Roques, S., Rebolo-Ifrán, N., & Tella, J. L. (2014). High urban breeding densities do not disrupt genetic monogamy in a bird species. *PLoS One*, *9*(3), e91314. https://doi.org/10.1371/journal.pone.0091314

Sánchez-Tójar, A., Parejo, D., Martínez, J. G., Rodríguezruiz, J., & Avilés, J. M. (2015). Parentage analyses reveal hidden breeding strategies of European rollers (*Coracias garrulus*). *Acta Ornithologica*, *50*(2), 252–258. https://doi.org/10.3161/00016454AO2015.50.2.013.

Schielzeth, H., & Bolund, E. (2010). Patterns of conspecific brood parasitism in zebra finches. *Animal Behaviour*, *79*(6), 1329–1337. https://doi.org/10.1016/j.anbehav.2010.03.006

Semel, B., & Sherman, P. W. (2001). Intraspecific parasitism and nest-site competition in wood ducks. *Animal Behaviour*, *61*(4), 787–803. https://doi.org/10.1006/anbe.2000.1657

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, *69*(1), 82–90. https://doi.org/10.1139/z91-013

Stai, S. M., & Hughes, C. R. (2003). Characterization of microsatellite loci in wild and domestic Muscovy ducks (*Cairina moschata*). *Animal Genetics*, *34*, 387–389.

Šťovíček, O., Kreisinger, J., Javůrková, V., & Albrecht, T. (2013). High rates of conspecific brood parasitism revealed by microsatellite analysis in a diving duck, the common pochard (*Aythya ferina*). *Journal of Avian Biology*, *44*(4), 369–375. https://doi.org/10.1111/j.1600-048X.2013.00074.x

Thow, C. M., Eadie, J. M., Wells, C. P., & Lyon, B. E. (2021a). Simulated wood duck maternity analysis results from COLONY and CERVUS. *Dryad*. https://doi.org/10.7291/D1VH5F

Thow, C. M., Eadie, J. M., Wells, C. P., & Lyon, B. E. (2021b). Simulated wood duck genotypes. *Dryad*. https://doi.org/10.7291/D1067B

Tiedemann, R., Paulus, K. B., Havenstein, K., Thorstensen, S., Petersen, A., Lyngs, P., & Milinkovitch, M. C. (2011). Alien eggs in duck nests: brood parasitism or a help from grandma? *Molecular Ecology*, *20*(15), 3237–3250. https://doi.org/10.1111/j.1365-294X.2011.05158.x

Toonen, R. J., & Hughes, S. (2001). Increased throughput for fragment analysis on ABI Prism 377 Automated Sequencer using a membrane comb and STRand software. *BioTechniques*, *31*, 1320–1324.

Tucker, A. M., Dyer, R. J., Huber, S. K., & Bulluck, L. P. (2016). Opportunistic conspecific brood parasitism in a box-nesting population of Prothonotary warblers (*Protonotaria citrea*). *The Auk: Ornithological Advances*, *133*(2), 298–307. https://doi.org/10.1642/AUK-15-161.1

Walling, C. A., Pemberton, J. M., Hadfield, J. D., & Kruuk, L. E. B. (2010). Comparing parentage inference software: Reanalysis of a red deer pedigree. *Molecular Ecology*, *19*(9), 1914–1928. https://doi.org/10.1111/j.1365-294X.2010.04604.x

Wang, J. (2016). *Users guide for software Colony version 2.0.6.2*.

Weinman, L. R., Solomon, J. W., & Rubenstein, D. R. (2015). A comparison of single nucleotide polymorphism and microsatellite markers for analysis of parentage and kinship in a cooperatively breeding bird. *Molecular Ecology*, *15*, 502–511. https://doi.org/10.1111/1755-0998.12330

Westneat, D. F., & Stewart, I. R. K. (2003). Extra-pair paternity in birds: Causes, correlates, and conflict. *Annual Review of Ecology, Evolution, and Systematics*, *34*, 365–396. https://doi.org/10.1146/annurev.ecolsys.34.011802.132439

Wiebe, K. L., & Kempenaers, B. (2009). The social and genetic mating system in flickers linked to partially reversed sex roles. *Behavioral Ecology*, *20*(2), 453–458. https://doi.org/10.1093/beheco/arn138

Yom-Tov, Y. (1980). Intraspecific nest parasitism in birds. *Biological Reviews*, *55*(1), 93–108. https://doi.org/10.1111/j.1469-185X.1980.tb00689.x

Yom-Tov, Y. (2001). An updated list and some comments on the occurrence of intraspecific nest parasitism in birds. *Ibis*, *143*(1), 133–143. https://doi.org/10.1111/j.1474-919X.2001.tb04177.x

Yom-Tov, Y., & Geffen, E. (2017). Conspecific brood parasitism among birds: The effects of phylogeny, mode of reproduction and geographic distribution. In M. Soler (Ed.), *Avian brood parasitism – behaviour, ecology, evolution and coevolution* (pp. 95–103). Springer.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Thow, C. M., Eadie, J. M., Wells, C. P., & Lyon, B. E. (2022). Pedigree simulations reveal that maternity assignment is reliable in populations with conspecific brood parasitism, incomplete parental sampling and kin structure. *Molecular Ecology Resources*, 22, 180–198. https://doi.org/10.1111/1755-0998.13466