Marine Mammal Science

ARTICLE OPEN ACCESS

Pedigree-Informed Abundance Estimation for the Endangered North Atlantic Right Whale (*Eubalaena glacialis*) via Gametic Mark-Recapture

Timothy R. Frasier[1] | Brenna A. Frasier[1] | Sonya N. Radvan[1] | Moira W. Brown[2] | Philip K. Hamilton[3] | Lisa A. Conger[4] | R. Clay George[5] | Katharine A. Jackson[6] | Scott D. Kraus[3]

[1]Department of Biology, Saint Mary's University, Halifax, Nova Scotia, Canada | [2]Canadian Whale Institute, Welshpool, New Brunswick, Canada | [3]Anderson Cabot Center for Ocean Life, New England Aquarium, Central Wharf, Boston, Massachusetts, USA | [4]Northeast Fisheries Science Center, NOAA-NMFS, Woods Hole, Massachusetts, USA | [5]Georgia Department of Natural Resources, Brunswick, Georgia, USA | [6]Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, St. Petersburg, Florida, USA

Correspondence: Timothy R. Frasier (timothy.frasier@smu.ca)

Received: 10 September 2025 | **Revised:** 6 December 2025 | **Accepted:** 17 December 2025

Keywords: abundance estimation | *Eubalaena glacialis* | gametic mark-recapture | genetics | pedigree | photo-identification | right whale

ABSTRACT

Pedigrees can provide important data for abundance estimation by providing a means to infer the presence of individuals that might otherwise not be detected. Such information may be particularly important for endangered species, where the addition or removal of just a few individuals can make the difference between population growth and decline. Here, we use gametic mark-recapture methods to incorporate pedigree information into male abundance estimates for the endangered North Atlantic right whale (*Eubalaena glacialis*)—a species for which there is conflicting information: some data suggest that almost all individuals have been identified by photo-identification efforts and are therefore accounted for, whereas other data suggest that a substantial number of individuals may not be identified and are therefore "missing" from abundance estimates. Our pedigree analyses did not infer a large number of North Atlantic right whales that were unaccounted for in the photo-identification data, suggesting that the photo-identification catalog is representative of the species as a whole and that abundance estimates based on those data are accurate.

1 | Introduction

Accurate abundance estimation is crucial for the management and conservation of species. Abundance estimates are an important consideration when listing and de-listing threatened and endangered species, and for determining the conservation priorities for listed species. Moreover, accurate abundance estimates are required to determine trends over time, such as identifying if a population is increasing or decreasing. Reliably detecting trends is key for evaluating the effectiveness of conservation plans and for developing new management or conservation strategies (e.g., Caughley and Gunn 1996).

There are a variety of methods to estimate abundance (summarized in Krebs 1989; Buckland et al. 2001). For species where individuals can be recognized, based either on naturally occurring physical characteristics or on artificial "marks," capture-mark-recapture (CMR) methods are often used (e.g., Krebs 1989; Amstrup et al. 2005; Kéry and Schaub 2012). Although CMR methods have become the cornerstone of abundance estimation for many populations, there remain some for which the effectiveness of this approach is limited. This is particularly true for populations that have a wide distribution that makes it difficult to "capture" an adequate representation of the population. For example, Gore et al. (2016) used photo-identification

and CMR methods to estimate the abundance of basking sharks (*Cetorhinus maximus*) in the Northeast Atlantic. However, due to limited recaptures resulting from the wide distribution of the population, it was not possible to obtain accurate abundance estimates.

For studies where high-resolution genetic data are combined with individual-based field data, pedigree information can be used to infer the presence of individuals that have not been "marked" and therefore improve abundance estimation (Jones and Avise 1997a, 1997b; Pearse et al. 2001; Israel and May 2010). As an early example, (Pearse et al. 2001) conducted a CMR study of painted turtles (*Chrysemys picta*) using genetic profiles as individual marks. By genetically sampling some individuals within the population, including offspring with one known parent, the second parent could be genetically inferred. This information was then used to estimate the effective number of breeders.

Originally, such studies were limited to estimating the number of breeders but have since evolved to being able to estimate total abundance—even in situations with limited field data—in approaches that have been termed "gametic mark-recapture" and "close-kin mark-recapture" (CKMR) (Skaug 2001; Bravington, Grewe, and Davies 2016; Ruzzante et al. 2019). Such approaches greatly improve the power to estimate abundance and monitor species and may be particularly helpful in cases where individual-based recognition is difficult in the field (e.g., Bravington, Skaug, and Anderson 2016; Hillary et al. 2018).

The endangered North Atlantic right whale (*Eubalaena glacialis*) represents one species (of which there is just one population) where pedigree data may be particularly useful for improving abundance estimation. This species represents one of the first and most extensively commercially exploited whale species (Aguilar 1986; Reeves et al. 2007) and has shown little signs of recovery since gaining international protection in 1935 (e.g., Pace III et al. 2017). Moreover, the species has been in decline since 2011 (Pace III et al. 2017). Limited recovery has been attributed to two main causes: (1) a high rate of anthropogenic mortality from vessel strikes and entanglement in commercial fishing gear (Kraus et al. 2005; Knowlton et al. 2012; Corkeron et al. 2018) and (2) a reproductive rate that is three times lower than their predicted potential (Browning et al. 2010; Frasier et al. 2024).

Individual right whales can be identified by the unique, naturally occurring callosity patterns on the top and sides of their heads, which remain stable throughout the life of each whale (Payne et al. 1983; Kraus et al. 1986). Photo-identification based on these callosity patterns and anthropogenic scarring is the primary source of population data on this species (e.g., Hamilton et al. 2007), and is also the source of data for CMR abundance estimation (e.g., Pace III et al. 2017; Reed et al. 2022; Linden 2024).

Dedicated photo-identification studies of North Atlantic right whales began in 1980. Consistent photo-identification surveys have been conducted annually in all major habitats in the Southeast U.S., Northeast U.S., and Atlantic Canada since 1994. As a result, the majority of whales that are seen each year already exist in the photo-identification catalog, and most sightings of "new" whales can be attributed to known births. Moreover, in most years over 90% of known individuals are sighted (Pettis et al. 2023). The apparently closed population and high recapture rates have led some researchers to regard the photo-identification catalog as a near census of the species (e.g., Clapham et al. 1999).

There are, however, three lines of evidence that suggest whales may exist that have never been photo-identified. First, although four critical habitats are currently recognized in the western North Atlantic (Fisheries and Oceans Canada 2014; National Marine Fisheries Service 2016), right whales have also been detected visually and acoustically in the central and eastern North Atlantic (Jacobsen et al. 2004; Silva et al. 2012; Davis et al. 2017). Although all but one of these whales have been matched to the photo-identification catalog, these sightings raise the possibility of other, undetected whales elsewhere in the North Atlantic. Second, there are whales that are sighted so infrequently (with many years passing in between sighting events) that they are referred to as "irregular" whales (Hamilton et al. 2007). The fact that these whales go multiple years without being sighted points to the likelihood of unidentified habitats and raises the question of how many whales *only* use these unidentified habitats and are therefore unknown to researchers. Lastly, due to long-term collaborations among many research teams, skin samples for genetic analyses have been collected for ~75% of all photo-identified individuals. These samples are regularly used to conduct paternity analyses and improve our understanding of reproduction within this species (Frasier et al. 2007, 2013; Crossman et al. 2024). Analyses of samples collected from 1980 to 2001 showed that the fathers had not been sampled for a surprisingly large percentage of the calves (49%), despite samples being available for the majority (69%) of photo-identified males (Frasier et al. 2007). The implication was that there may be males within the species which are not accounted for based solely on the photo-identification data.

Taken together, these data provide conflicting information. Based solely on the photo-identification data, it appears that most whales have been identified, and therefore that the catalog is nearly representative of all individuals remaining in the species. On the other hand, multiple lines of evidence suggest that there may be individuals that do not use identified habitats, are not "captured" via photo-identification, have different capture probabilities than most other individuals, and may therefore not be appropriately accounted for in analyses or interpretations of this species to date. Most adult females are dependably sighted (and can be sampled) in the calving grounds in the Southeast U.S. in years that they calve. However, there is no known habitat that aggregates the majority of males similarly in space or time. It therefore seems likely that adult males may be more apt to be "missing" from the photo-identification and genetics databases, as suggested by earlier paternity analyses.

Here, we combine long-term photo-identification data with high-resolution genetic data to conduct parentage analyses and calculate pedigree-informed abundance estimates for males of the endangered North Atlantic right whale. Our methods follow those developed for such analyses ("gametic

mark-recapture"), and which have been used in similar contexts for other whale species (Garrigue et al. 2004; Carroll et al. 2012; Førland et al. 2026). The goal was to identify if the genetic inference of individuals results in higher abundance estimates than those based solely on photo-identification data and therefore improve such estimates for this endangered species.

2 | Methods

2.1 | Photo-Identification Data

Individual North Atlantic right whales can be identified photographically using the callosity patterns on their heads (Payne et al. 1983; Kraus et al. 1986), as well as by variation in pigmentation patterns and scars elsewhere on their body (Kraus 1990). Consistent photo-identification surveys of key habitats began in 1980 and have continued to the present. Photographs for photo-identification are collected by many research teams and are sent to the New England Aquarium right whale research team, which conducts photo-identification analyses and curates and maintains the photo-identification database and catalog. The photo-identification data used for the paternity analyses in this study were obtained from the North Atlantic Right Whale Consortium (NARWC), which is a collaboration of researchers, government, and non-governmental organizations formed to share data, information, and ideas with respect to North Atlantic right whales. We obtained photo-identification data from 1980 to 2021 (North Atlantic Right Whale Consortium 2023).

2.2 | Genetic Data

Small tissue samples have been collected from free-swimming North Atlantic right whales, using a crossbow with a modified bolt, since 1988 (Brown et al. 1991). This is the primary means by which samples are collected from free-swimming whales around the world. Intensive and long-term studies have shown that this method does not have negative short- or long-term effects on the whales, other than an initial startle response, seen a small proportion of the time (e.g., Brown et al. 1991; Best et al. 2005; Noren and Mocklin 2011). Samples are collected by the same research groups that also collect the photo-identification data, and photo-identification data are collected in association with each sampling event. Routine genetic analyses conducted on each sample include genetic sex-determination (Shaw et al. 2003; Gilson et al. 1998; Konrad et al. 2017), sequence analysis of a portion of the mitochondrial control region (Malik et al. 1999), and genotype analysis at 35 microsatellite loci (Frasier et al. 2006, 2007). Although all samples are genotyped at 35 loci (using loci and protocols described in Frasier et al. 2006), 28 of these are used for subsequent analyses. Two loci are removed from analyses due to high estimates of null alleles and five are removed due to significant signs of linkage (for details see Frasier et al. 2007). Individuals are considered "genotyped" and used in paternity analyses if they have data for at least 22 of these 28 loci. A graph showing the percentage of individuals genotyped at each number of loci is shown in Figure S3.

2.3 | Paternity Analyses

The photo-identification and genetic data were combined to conduct paternity analyses for each year. Because right whale calves suckle for 10–12 months and maintain a close association with their mothers during that time (Hamilton et al. 1998), maternity can be determined photographically and genetically by collecting and linking images and genetic sampling results from mothers and calves during the calves' first year of life. There has been an ongoing effort to collect genetic samples from as many reproductive age females and calves as possible, and the resulting calving and maternity data are maintained in the photo-identification database. For each year, maternity was confirmed by using genotypes of sampled mother-calf pairs (e.g., Frasier et al. 2010), and paternity was assigned by comparing these mother-calf genotypes with those from all whales considered "candidates" (Frasier et al. 2007). Briefly, whales were considered as candidates for paternity analyses in a given year if they fell into one of three categories:

1. **Known-age candidates:** Known-age males or individuals of unknown sex that were at least 5 years old in the year of the calf's conception (the year prior to when the calf was born). We do not know at what age males can first reproduce; however, the average age at first conception for females is 10 years (Kraus et al. 2007), but there is one female that became pregnant at the age of 4 years (Knowlton et al. 1994). We do not want to incorrectly exclude younger males from paternity analyses, and therefore we conservatively chose an age cut-off of 5 years. Previous work has shown that paternities are biased toward older males, with the youngest assigned father for the years between 1980 and 2001 being 15 years old (Frasier et al. 2007). For sexes, at the time of analyses the photo-identification database had 334 females, 388 males, and 77 individuals of unknown sex; and the genetic database contained 301 females and 358 males.
2. **Unknown-age candidates 1:** Unknown-age males and individuals of unknown sex sighted in years prior to when the calf was born and that were still presumed to be alive in the year of conception. Because over 90% of individual right whales are sighted annually (Pettis et al. 2023), whales that have not been sighted for six consecutive years are classified as "presumed dead" (Knowlton et al. 1994). These candidates represent whales that were likely alive and adult in the year of conception and should therefore be considered as candidates, but that could have died in the years after conception.
3. **Unknown-age candidates 2:** Unknown-age males and individuals of unknown sex not sighted prior to when the calf was born, but sighted no more than 5 years after. These represent whales that may have been male, sexually mature, and alive in the year of conception, but did not have their first sighting until after the year the calf was born. This 5-year window should be conservative (i.e., gives males ample time to be sighted) because in most years over 90% of known individuals are sighted (Pettis et al. 2023).

The need for criteria two and three has decreased throughout the study period and will continue to do so. Early in the study

period (e.g., in the 1980s), most whales were added to the photo-identification catalog when they were juveniles or adults and were of unknown age. In recent decades, however, an increasing percentage of whales in the catalog were first observed during their birth year and are therefore of known age.

In the past, we used the program CERVUS (Marshall et al. 1998; Kalinowski et al. 2007) to identify potential mother-calf-father triads, and paternity was assigned based on exclusion (i.e., in cases where only one male remained non-excluded as the father based on genetic mismatches). We chose this approach because the genotyping error rate in our data set is low (Frasier et al. 2009), and we did not think that the statistical approaches used by the CERVUS software for assigning confidence in parentage assignments were biologically appropriate. However, we have since changed our approach.

First, as our data set has grown we recognize that errors are unavoidable, but we want to minimize the effects of any errors on paternity assignment. The exclusion method is intolerant of mismatches at even a single locus, and is therefore highly impacted by even small errors (e.g., Flanagan and Jones 2019). Moreover, new approaches have been developed for calculating the statistical confidence of paternity assignments, such as those implemented in the *MasterBayes* R package (Hadfield et al. 2006; Koch et al. 2008), in ways that are more biologically appropriate and result in more accurate assignments (Christie 2013; Christie et al. 2013; Hadfield et al. 2006; Koch et al. 2008). For example, Walling et al. (2010) compared several paternity assignment approaches in the red deer (*Cervus elaphus*) population on the Isle of Rum, Scotland, and found that the approaches implemented in *MasterBayes* resulted in more correct paternity assignments and fewer incorrect assignments than obtained using similar criteria implemented in CERVUS. The main differences are: (1) that *MasterBayes* can include phenotypic information in paternity assessment, such as whether or not a female was in a particular male's harem at the time of conception; (2) *MasterBayes* can include estimating the size of the unsampled population in the analyses, and use that (and the uncertainty around this estimate) as part of the analyses, whereas in CERVUS the size of the unsampled population must be explicitly provided by the users, and which can have large impacts on confidence in paternity assignments (e.g., Koch et al. 2008); and (3) confidence in paternity is assessed at the individual level in *MasterBayes*, whereas it is calculated at the population level in CERVUS.

With respect to difference (3), CERVUS conducts simulations to calculate confidence in paternity assignments; generating simulated parents, offspring, and other candidate males from the allele frequencies in the study population. It then samples this synthetic population in the same manner as the study population (i.e., using the user-defined percentage of sampled candidate males, which may or may not include the true fathers), and calculates the differences in the log-likelihood of paternity for the two most likely males (called "delta") for each synthetic offspring. It compares this delta value from the observed candidates to the simulated data set, which are used to identify cut-off values for different confidence levels. For example, for 95% confidence, CERVUS identifies at what delta value were 95% of the synthetic paternity assignments correct. Any delta value greater than or equal to this value in the study data set is then assigned at the 95% confidence level. This approach means that confidence in paternity assignment in CERVUS is calculated at the *population level* rather than at the individual level, because the critical values for paternity assignment at a given confidence level are based on the genetic characteristics of the population (via the simulations), rather than any specific potential mother-calf-father triad. The approach used in *MasterBayes*, however, calculates the posterior probability for each paternity assignment that users can then filter based on their cut-off of choice. In this sense, confidence in paternity assignments is based on each case itself (i.e., at the individual level) rather than based on comparison to a population-level cut-off value. Because of this approach, using a 95% population-level criterion in CERVUS can result in paternities being assigned to males with individual probabilities of paternity being much lower than 95%. Conversely, a lower criterion in *MasterBayes* (e.g., 85%) may actually be more stringent (less likely to have an incorrect male assigned paternity) than using a 95% criterion in CERVUS (e.g., Walling et al. 2010).

For these reasons we have shifted to using *MasterBayes* v.2.58 for paternity assignment. We did so on a year-by-year basis, with different mother-calf pairs and a different pool of candidate males for each year. Briefly, for each year we: (1) identified the known mother-calf pairs, and those for which genotypes had been obtained; (2) confirmed these maternities based on the genetic data; (3) identified the appropriate suite of "candidate" males for that year based on the criteria described above; and (4) conducted paternity analysis using *MasterBayes*. For these analyses, preliminary trials indicated that there was not enough information in the data to estimate genotyping error rates. We therefore used a pre-defined value of 0.005 for both error rates: the rate of allelic dropout (E1) and the rate of other discrepancies (mutations and genotyping errors, E2). We ran the analyses using 5000 MCMC steps as burn-in and 20,000 steps with recorded data. We considered paternity "assigned" for those paternities with a posterior probability greater than 0.95. For comparison purposes, we also conducted paternity analyses for each year using CERVUS v.3.07 using paternities assigned at the 80% and 95% confidence levels. It should be noted that *MasterBayes* has the functionality to estimate the size of the unsampled population, which is what we are trying to do and would therefore negate the need to perform gametic mark-recapture analyses. However, we were not able to obtain stable/reliable estimates in this manner, and therefore proceeded with the gametic-based mark-recapture analyses.

2.4 | Gametic Mark-Recapture

To use these pedigree data in gametic mark-recapture analyses, we used the equation below from Chapman (1951), as used by Garrigue et al. (2004) and Carroll et al. (2012) for similar purposes.

$$N_m = \frac{(n_1 + 1)(n_2 + 1)}{(m + 1)} - 1$$

where N_m is the estimated male abundance for that year, n_1 is the total number of genotyped candidates for that year, n_2 is the number of mother-calf pairs sampled in that year, and m is the

number of identified paternities. This is a modification of the Peterson method where the genotyped candidates represent the initial "marking" events, the genotyped mother-calf pairs are representative of the individuals subsequently "captured." and the paternity assignments represent those candidates that were subsequently "recaptured."

We used a Bayesian approach to obtain estimates of N_m and the associated uncertainty for each year. To do this, we estimated the probability of a male being "captured" as a father (Θ), given the sample size (number of genotyped mother-calf pairs) for each year using a binomial model. We did this in a hierarchical manner, where the estimates from each year informed those estimates across all years. Specifically, for each of y years, the equation used was:

$$(m_y + 1) \sim \text{binomial}\left(n_{2y} + 1, \theta_y\right)$$

where the prior used for the Θ was a normal distribution with the hierarchical priors Θ_{mean} and Θ_{sd}. The prior for Θ_{mean} was a uniform distribution between 0 and 1, and the prior for Θ_{sd} was an exponential distribution with a rate of 0.5.

The values of $n_1 + 1$ for each year were then divided by Θ for each year, and then 1 was subtracted to obtain annual estimates of N_m.

Data and code are available at https://github.com/timothyfrasier/pedigreeAbundance. These analyses were conducted and visualized using a combination of R v.4.4.2 (R Core Team 2024), RStudio v.2024.12.0.467 (Posit Team 2024), Stan v.2.36 (Stan Development Team 2024), cmdstanr v.0.8.1 (Gabry et al. 2024), and ggplot2 v.3.5.1 (Wickham 2016).

3 | Results and Discussion

3.1 | Paternity Assignment

Across the 44 years from 1980 to 2023, a total of 315 mother-calf pairs has been sampled and genotyped and therefore available for paternity assignment (Table 1, Figure S1). This represents ~51% of the 612 known mother-calf pairs between 1980 and 2023. The total number of paternities assigned by the different methods was: Exclusion—195, CERVUS with 80% criterion—235, CERVUS with 95% criterion—206, and *MasterBayes* with 95% criterion—243 (Table 1, Figure S2).

Overall, there was substantial agreement among paternity assignments across methods (Figure 1, Table 1). Most notably, there were not cases where a different male was assigned paternity using different methods: all differences were due to one method assigning paternities that another did not. The biggest differences were those paternities not assigned by exclusion, but which were assigned by one or more of the probabilistic-based methods; particularly the 30 not assigned by exclusion but assigned by all other methods/criteria (Figure 1). This result (and the associated details) indicates that using an exclusion-based paternity assignment approach—which does not allow for genotyping errors—will result in incorrect mismatches and therefore incorrect exclusion of true paternities. The result is fewer paternity assignments than when using probabilistic-based methods, even in data sets with low error rates, as predicted by Flanagan and Jones (2019). It is also noteworthy that there were three paternities assigned by exclusion that were not assigned by any other methods (Figure 1). These represent males that did not have any mismatches, but that were still not assigned at a high probability by any other method, presumably due to them possessing very common alleles and/or the presence of multiple loci where calves were heterozygous for alleles present in both putative parents—leading to ambiguities in which alleles were inherited from the mother versus the father.

Similar to other studies, such as that of red deer (Walling et al. 2010), we found that the approach used by *MasterBayes* resulted in a similar number of paternity assignments as using the 80% criterion in CERVUS, but with *MasterBayes* assigning them at a much higher probability (i.e., ≥95%). This has been attributed to confidence levels being calculated at the population versus individual level, respectively (Walling et al. 2010). The method for estimating critical cut-off values for different confidence levels in CERVUS involves using the observed allele frequencies to generate simulated individuals, creating synthetic offspring from those individuals, calculating the difference in likelihood between the two most likely fathers (called "delta"), and then identifying at which delta value the desired level of confidence is obtained (e.g., for the 95% confidence cut-off value, identifying at which delta value are 95% of the assigned paternities correct). Thus, this cut-off value is akin to an average across all assignments for a given level of confidence, and means that many true fathers will not meet the chosen population-wide criterion (Type II error), and that some incorrect paternity assignments will "pass" the criterion (Type I error). *MasterBayes*, on the other hand, bases assignments on the probability for each particular putative mother-offspring-father triad. This approach should reduce the rate of both Type I and Type II errors. Type II errors are arguably more common in paternity assignment studies when a large number of markers are used and stringent criteria applied. Therefore, the reduction of Type II errors using the approach implemented by *MasterBayes* results in a larger number of assignments at higher levels of confidence.

3.2 | Abundance Estimation

Results from the gametic mark-recapture analyses are shown in Figure 2. For the most part, the number of males estimated via gametic mark-recapture aligns very closely with those based on photo-identification, suggesting that there are not a substantial number of males "missing" from the photo-identification catalog. There are, however, a few years where the estimates differ. Specifically, there are a series of years (2006–2009) where the pedigree-informed estimates are less than those from photo-identification, and a series of more recent years (2019–2021) where this pattern is reversed. It is not clear what is resulting in the lower estimates for the years 2006–2009. Theoretically, this could result in years where a particularly high number of calves have paternities assigned, resulting in a lower number of inferred males. This appears to be the case, with the years 2005–2009 having a higher number

TABLE 1 | Summary of paternity analysis data and results.

Year	Candidates	M-C pairs	Exclusion	C. 80%	C. 95%	MB 95%	Inferred	Lower	Upper
1980	71	0	0	0	0	0	90	79	102
1981	78	3	2	2	0	2	99	88	113
1982	80	6	5	5	5	5	101	90	113
1983	87	3	1	2	1	2	110	99	126
1984	89	8	4	4	4	5	114	103	131
1985	92	2	2	2	2	2	115	102	130
1986	92	6	4	4	4	4	117	106	134
1987	98	7	5	5	4	5	124	112	142
1988	103	3	2	2	2	3	129	113	145
1989	105	12	8	9	8	9	133	121	151
1990	116	7	3	3	2	3	150	137	177
1991	118	6	4	4	4	4	150	136	171
1992	125	6	2	3	2	3	161	146	187
1993	131	2	1	1	1	1	166	150	191
1994	134	5	4	4	3	4	169	151	191
1995	142	3	3	3	3	3	177	156	200
1996	142	10	7	8	8	9	177	158	197
1997	149	9	8	8	8	8	186	166	207
1998	154	1	0	0	0	0	196	176	226
1999	149	2	1	2	1	2	187	165	211
2000	152	0	0	0	0	0	191	170	218
2001	153	18	9	13	11	13	195	178	221
2002	161	11	5	9	7	8	205	186	232
2003	163	15	8	10	9	10	209	191	240
2004	170	12	9	10	8	11	211	187	233
2005	164	19	14	15	13	16	205	185	228
2006	164	14	11	13	12	13	202	179	224
2007	171	17	11	15	13	15	212	190	235
2008	173	19	14	17	16	18	212	186	232
2009	183	28	14	19	16	20	235	216	267
2010	192	13	6	6	6	8	248	226	288
2011	191	8	7	7	7	7	239	213	268
2012	195	1	0	0	0	0	248	223	286
2013	199	9	6	6	6	7	251	226	283
2014	211	2	1	2	2	2	264	234	298
2015	220	11	5	9	8	9	276	248	310
2016	226	7	3	5	4	5	287	259	327
2017	222	1	0	0	0	0	282	254	326

(Continues)

TABLE 1 | (Continued)

Year	Candidates	M-C pairs	Exclusion	C. 80%	C. 95%	MB 95%	Inferred	Lower	Upper
2018	207	0	0	0	0	0	261	231	298
2019	206	1	1	1	1	1	259	229	293
2020	201	7	4	6	4	6	252	226	282
2021	199	1	1	1	1	0	253	228	292
2022	189	0	0	0	0	0	238	211	272
2023	177	0	0	0	0	0	223	198	253
Total		315	195	235	206	243			

Note: Included are: the year; the number genotyped candidates for paternities; the number of genotyped mother-calf pairs; the number of paternities assigned based on exclusion, the 80% criterion with CERVUS, the 95% criterion with CERVUS, and the 95% criterion with *MasterBayes*; and the mean number of inferred males as well as the lower and upper 95% highest density intervals (HDI).

of paternity assignments than the other years throughout the study (Figure S2). The difference in more recent years (2019–2021), which coincides with a time period when photo-identification-based estimates consistently show a concerning decline in abundance (Pettis et al. 2023; Linden 2024), may at first seem surprising. However, these results should *not* be interpreted as refuting that decline. Instead, Table 1 shows that we have very few sampled mother-calf pairs for the latter years of this study (due to the time required to obtain samples from specific individuals); specifically, just 1 pair for 2017, 0 for 2018 (when no calves were observed), 1 for 2019, 7 for 2020, 1 for 2021, 0 for 2022, and 0 for 2023. Our genetic-based estimates for these latter years are therefore based on very little data, and should be interpreted with caution rather than as conflicting with the photo-identification-based estimates. Indeed, the requirement for additional data (photo-identification data for mother-calf associations, then samples and genotypes for mother-calf pairs and candidate fathers) should be viewed as a limitation of this gametic-based approach. Regardless, despite these occasional differences, the results show strong agreement between the photo-identification and pedigree-informed estimates.

This result, suggesting that there are not a large number of "missing" males, was somewhat surprising given the results from previous paternity analyses. Frasier et al. (2007) reported on paternity analyses of 87 mother-calf pairs conducted in 2002, for the years 1980–2001, where genetic profiles were available for 69% of all identified candidate males but paternity was only assigned for 51% of the calves. The implication was that this may be evidence for the existence of males that were not part of the photo-identification catalog. There are likely several factors resulting in the difference between the results from the previous study and the results presented here. The biggest factor is likely sample size and representation. Our current sample set is much larger: we conducted paternity analyses on 315 genotyped mother-calf pairs from 1980 to 2023. Thus, we have added many more sampled mother-calf pairs to the paternity analyses. Additionally, we currently have a much higher percentage of candidate males genotyped (an average of about 75% per year). Combined, these have resulted in better agreement between the observed and expected number of paternity assignments: paternity was assigned to 243 of the 315 genotyped calves (77%), which is about what would be expected given the percentage of genotyped candidate males. Another reason is likely due to paternity assignment method. The original study used exclusion for paternity assignment, which—as shown above—can artificially reduce the number of paternities assigned due to mismatches caused by genotyping errors. Our method of using the 95% criterion with *MasterBayes* resulted in many more paternities assigned than with exclusion (243 vs. 195), presumably due to a great reduction in falsely rejected paternity assignments (Type II errors). Therefore, this increase in sample size and representation, as well as the implementation of a different paternity assignment method that allows for genotyping errors, likely represent the reasons why the results presented here differ from those reported in earlier studies, with respect to the number of calves for which paternity was not assigned.

Our results do not support the inference of a substantial number of "missing" males, but rather suggest that the cataloged population is representative of the entire species. This finding has important implications for conservation. With the results of the previous paternity study, there was a hopeful notion of some number of "unknown" whales swimming in less urban and degraded habitats elsewhere in the North Atlantic Ocean, with the implication that they could help maintain or even replenish the population as more urban areas deteriorate. However, our current results suggest that this idea is unlikely, and that the whales we photograph and catalog represent the vast majority of the species. This increases the importance and urgency of protecting the cataloged whales and their habitats, because there are no others that can come to their rescue. These results also reiterate the value of the photo-identification and monitoring efforts, which combine to provide data for essentially all individuals within the species.

More broadly, our results add to the growing literature on the application of gametic mark-recapture methods in conservation biology. On one hand, they highlight that the power of these methods is limited in cases where few mother-offspring pairs (or other "reference" individuals) are sampled within each analyzed time period, specifically in cases where there is a wealth of other identification (e.g., photo-identification) data available. However, these results also provide a means to infer individuals not detected by other methods, which is extremely valuable

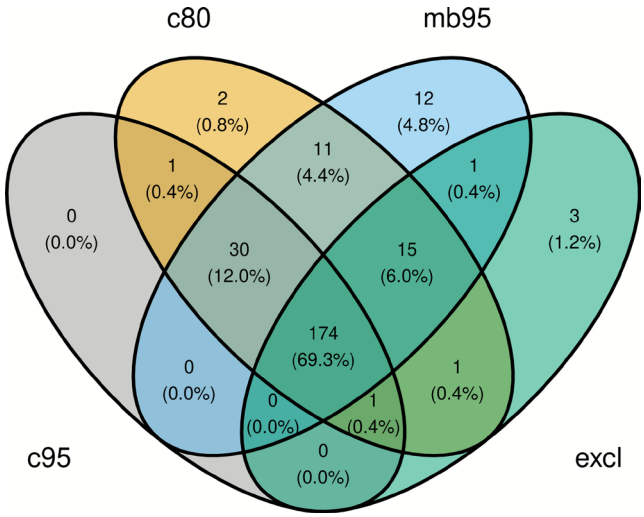

FIGURE 1 | Venn diagram of paternity assignments using the four different methods/criteria. "c80" and "c95" refer to paternities assigned by the program CERVUS using the 80% and 95% confidence criterion, respectively. "excl" and "mb95" refer to paternities assigned using exclusion and *MasterBayes* with a 95% probability threshold, respectively.

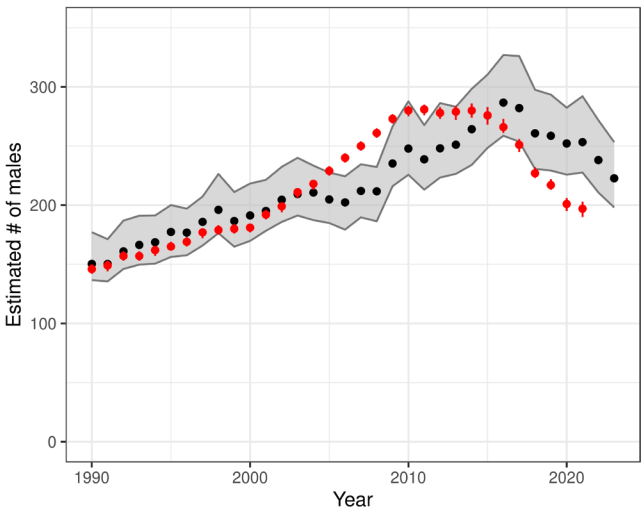

FIGURE 2 | Abundance estimates for the number of males based on gametic mark-recapture (black dots, gray ribbon represents 95% highest density intervals [HDI]), and based on photo-identification (red dots and error bars). The photo-identification abundance estimates are from the model described in Pace III et al. (2017) as presented in Frasier et al. (2024). Note that the photo-identification estimates are for all males, whereas the gametic mark-recapture estimates are for adult males.

and informative. Indeed, rejecting our hypothesis of a significant number of "missing" males is highly valuable: it answers a major question about this species and allows us to move forward with the photo-identification-based estimates with confidence. However, for many other species—particularly those where individual identification via other means is more difficult—this approach represents a powerful tool for population assessment and monitoring (e.g., Bravington, Skaug, and Anderson 2016; Hillary et al. 2018).

Author Contributions

Timothy R. Frasier: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, visualization, writing – original draft.

Acknowledgments

The routine genetic analyses of North Atlantic right whale samples and the maintenance of the North Atlantic right whale genetic database are currently funded by Fisheries and Oceans Canada (Contract to T.R.F.). This work has been ongoing for over 30 years, and we are indebted to Dr. Bradley N. White who developed and maintained this database for decades. We would also like to thank all people and teams who have collected samples for genetic analyses, and all of the many people and teams that have collected photo-identification data since 1980 and contributed them to the North Atlantic Right Whale Consortium. We also thank Richard M. Pace III for sharing data on the estimated number of adult males alive for each year.

Funding

This work was supported by Fisheries and Oceans Canada.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in github at https://github.com/timothyfrasier/pedigreeAbundance.

References

Aguilar, A. 1986. "A Review of Old Basque Whaling and Its Effect on the Right Whales (*Eubalaena glacialis*) of the North Atlantic." *Report of the International Whaling Commission Special Issue* 10: 191–199.

Amstrup, S. C., T. L. McDonald, and B. F. J. Manly. 2005. *Handbook of Capture-Recapture Analysis*. Princeton University Press.

Best, P. B., D. Reeb, M. B. Rew, P. J. Palsbøll, C. M. Schaeff, and A. Brandão. 2005. "Biopsying Southern Right Whales: Their Reactions and Effects on Reproduction." *Journal of Wildlife Management* 69: 1171–1180.

Bravington, M. V., P. M. Grewe, and C. R. Davies. 2016. "Absolute Abundance of Southern Bluefin Tuna Estimated by Close-Kin Mark-Recapture." *Nature Communications* 7: 13162.

Bravington, M. V., H. J. Skaug, and E. C. Anderson. 2016. "Close-Kin Mark-Recapture." *Statistical Science* 31: 259–274.

Brown, M. W., S. D. Kraus, and D. E. Gaskin. 1991. "Reaction of North Atlantic Right Whales (*Eubalaena glacialis*) to Skin Biopsy Sampling for Genetic and Pollutant Analysis." *Report of the International Whaling Commission Special Issue* 13: 81–90.

Browning, C. L., R. M. Rolland, and S. D. Kraus. 2010. "Estimated Calf and Perinatal Mortality in Western North Atlantic Right Whales (*Eubalaena glacialis*)." *Marine Mammal Science* 26: 648–662.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press.

Carroll, E. L., S. J. Childerhouse, M. Christie, et al. 2012. "Paternity Assignment and Demographic Closure in the New Zealand Southern Right Whale." *Molecular Ecology* 21: 3960–3973.

Caughley, G., and A. Gunn. 1996. *Conservation Biology in Theory and Practice*. Blackwell Science.

Chapman, D. G. 1951. "Some Properties of the Hypergeometric Distribution With Applications to Zoological Censuses." *University of California Publications in Statistics* 1: 131–160.

Christie, M. R. 2013. "Bayesian Parentage Analysis Reliably Controls the Number of False Assignments in Natural Populations." *Molecular Ecology* 22: 5731–5737.

Christie, M. R., J. A. Tennessen, and M. S. Blouin. 2013. "Bayesian Parentage Analysis With Systematic Accountability of Genotyping Error, Missing Data and False Matching." *Bioinformatics* 29: 725–732.

Clapham, P. J., S. B. Young, and R. L. Brownell. 1999. "Baleen Whales: Conservation Issues and the Status of the Most Endangered Populations." *Mammal Review* 29: 37–62.

Corkeron, P., P. Hamilton, J. Bannister, et al. 2018. "The Recovery of North Atlantic Right Whales, *Eubalaena glacialis*, Has Been Constrained by Human-Caused Mortality." *Royal Society Open Science* 5: 180892.

Crossman, C. A., P. K. Hamilton, M. W. Brown, et al. 2024. "Effects of Inbreeding on Reproductive Success in Endangered North Atlantic Right Whales." *Royal Society Open Science* 11: 240490.

Davis, G. E., M. F. Baumgartner, J. M. Bonnell, et al. 2017. "Long-Term Passive Acoustic Recordings Track the Changing Distribution of North Atlantic Right Whales (*Eubalaena glacialis*) From 2004 to 2014." *Scientific Reports* 7: 13460.

Fisheries and Oceans Canada. 2014. *Recovery Strategy for the North Atlantic Right Whale (Eubalaena glacialis) in Atlantic Canadian Waters [Final]*. Fisheries; Oceans Canada.

Flanagan, S. P., and A. G. Jones. 2019. "The Future of Parentage Analysis: From Microsatellites to SNPs and Beyond." *Molecular Ecology* 28: 544–567.

Førland, B., H. J. Skuag, M. Takahashi, and L. A. Pastene. 2026. "Close-Kin Mark-Recapture Without Age Information: Application to Southern Right Whales." *Journal of Agricultural, Biological, and Environmental Statistics*. https://doi.org/10.1007/s13253-025-00689-3.

Frasier, T. R., R. M. Gillett, P. K. Hamilton, M. W. Brown, S. D. Kraus, and B. N. White. 2013. "Postcopulatory Selection for Dissimilar Gametes Maintains Heterozygosity in the Endangered North Atlantic Right Whale." *Ecology and Evolution* 3: 3483–3494.

Frasier, T. R., P. K. Hamilton, M. W. Brown, et al. 2007. "Patterns of Male Reproductive Success in a Highly Promiscuous Whale Species: The Endangered North Atlantic Right Whale." *Molecular Ecology* 16: 5277–5293.

Frasier, T. R., P. K. Hamilton, M. W. Brown, S. D. Kraus, and B. N. White. 2009. "Sources and Rates of Errors in Methods of Individual Identification for North Atlantic Right Whales." *Journal of Mammalogy* 90: 1246–1255.

Frasier, T. R., P. K. Hamilton, M. W. Brown, S. D. Kraus, and B. N. White. 2010. "Reciprocal Exchange and Subsequent Adoption of Calves by Two North Atlantic Right Whales (*Eubalaena glacialis*)." *Aquatic Mammals* 36: 115–120.

Frasier, T. R., P. K. Hamilton, and R. M. Pace III. 2024. "How Compromised Is Reproductive Performance in the Endangered North Atlantic Right Whale?" *Endangered Species Research* 55: 267–271.

Frasier, T. R., T. Rastogi, M. W. Brown, P. K. Hamilton, S. D. Kraus, and B. N. White. 2006. "Characterization of Tetranucleotide Microsatellite Loci and Development and Validation of Multiplex Reactions for the Study of Right Whale Species (Genus *Eubalaena*)." *Molecular Ecology Notes* 6: 1025–1029.

Gabry, J., R. Češnovar, A. Johnson, and S. Bronder. 2024. "Cmdstanr: R Interface to 'CmdStan'." https://mc-stan.org/cmdstanr/.

Garrigue, C., R. Dodemont, D. Steel, and C. S. Baker. 2004. "Organismal and 'Gametic' Capture-Recapture Using Microsatellite Genotyping Confirm Low Abundance and Reproductive Autonomy of Humpback Whales on the Wintering Grounds of New Caledonia." *Marine Ecology Progress Series* 274: 251–262.

Gilson, A., M. Syvanen, K. Levine, and J. Banks. 1998. "Deer Gender Determination by Polymerase Chain Reaction: Validation Study and Application to Tissues, Bloodstains, and Hair Forensic Samples From California." *California Fish & Game* 84: 159–169.

Gore, M. A., P. H. Frey, R. F. Ormond, H. Allan, and G. Gilkes. 2016. "Use of Photo-Identification and Mark-Recapture Methodology to Assess Basking Shark (*Cetorhinus maximus*) Populations." *PLoS One* 11: e0150160.

Hadfield, J. D., D. S. Richardson, and T. Burke. 2006. "Towards Unbiased Parentage Assignment: Combining Genetic, Behavioural and Spatial Data in a Bayesian Framework." *Molecular Ecology* 15: 3715–3730.

Hamilton, P. K., A. R. Knowlton, and M. K. Marx. 2007. "Right Whales Tell Their Own Stories: The Photo-Identification Catalog." In *The Urban Whale: North Atlantic Right Whales at the Corssroads*, edited by S. D. Kraus and R. M. Rolland, 75–104. Harvard University Press.

Hamilton, P. K., M. K. Marx, and S. D. Kraus. 1998. "Weaning in North Atlantic Right Whales." *Marine Mammal Science* 11: 386–390.

Hillary, R. M., M. V. Bravington, T. A. Patterson, et al. 2018. "Genetic Relatedness Reveals Total Population Size of White Sharks in Eastern Australia and New Zealand." *Scientific Reports* 8: 2661.

Israel, J. A., and B. May. 2010. "Indirect Genetic Estimates of Breeding Population Size in the Polyploid Green Sturgeon (*Acipenser medirostris*)." *Molecular Ecology* 19: 1058–1070.

Jacobsen, K. O., M. Marx, and N. ØIen. 2004. "Two-Way Trans-Atlantic Migration of a North Atlantic Right Whale (*Eubalaena glacialis*)." *Marine Mammal Science* 20: 161–166.

Jones, A. G., and J. C. Avise. 1997a. "Polygynandry in the Dusky Pipefish *Syngnathus floridae* Revealed by Microsatellite DNA Markers." *Evolution* 51: 1611–1622.

Jones, A. G., and J. C. Avise. 1997b. "Microsatellite Analysis of Maternity and the Mating System in the Gulf Pipefish *Syngnathus scovelli*, a Species With Male Pregnancy and Sex-Role Reversal." *Molecular Ecology* 6: 203–213.

Kalinowski, S. T., M. L. Taper, and T. C. Marshall. 2007. "Revising How the Computer Program Cervus Accommodates Genotyping Error Increases Success in Paternity Assignment." *Molecular Ecology* 16: 1099–1106.

Kéry, M., and M. Schaub. 2012. *Bayesian Population Analysis Using WinBUGS. A Hierarchical Perspective*. Academic Press.

Knowlton, A. R., P. K. Hamilton, M. K. Marx, H. M. Pettis, and S. D. Kraus. 2012. "Monitoring North Atlantic Right Whale *Eubalaena glacialis* Entanglement Rates: A 30 Yr. Retrospective." *Endangered Species Research* 466: 293–302.

Knowlton, A. R., S. D. Kraus, and R. D. Kenney. 1994. "Reproduction in North Atlantic Right Whales (*Eubalaena glacialis*)." *Canadian Journal of Zoology* 72: 1297–1305.

Koch, M., J. D. Hadfield, K. M. Sefc, and C. Sturmbauer. 2008. "Pedigree Reconstruction in Wild Cichlid Fish Populations." *Molecular Ecology* 17: 4500–4511.

Konrad, C. M., A. Dupuis, S. Gero, and T. Frasier. 2017. "A Sexing Technique for Highly Degraded Cetacean DNA." *Aquatic Mammals* 43: 655–660.

Kraus, S. D. 1990. "Rates and Potential Causes of Mortality in North Atlantic Right Whales (*Eubalaena glacialis*)." *Marine Mammal Science* 6: 278–291.

Kraus, S. D., M. W. Brown, H. Caswell, et al. 2005. "North Atlantic Right Whales in Crisis." *Science* 309: 561–562.

Kraus, S. D., K. E. Moore, C. A. Price, et al. 1986. "The Use of Photographs to Identify Individual North Atlantic Right Whales

(*Eubalaena glacialis*)." *Report of the International Whaling Commission Special Issue* 10: 145–151.

Kraus, S. D., Pace, R. M., III, and T. R. Frasier. 2007. "High Investment, Low Return: The Strange Case of Reproduction in *Eubalaena glacialis*." In *The Urban Whale: North Atlantic Right Whales in the Crossroads*, 172–199. Harvard University Press.

Krebs, C. J. 1989. *Ecological Methodology*. Harper Collins.

Linden, D. W. 2024. "Population Size Estimation of North Atlantic Right Whales From 1990–2023." U.S. Department of Commerce: Northeast Fisheries Science Center Technical Memorandum.

Malik, S., M. Brown, S. Kraus, and A. Knowlton. 1999. "Assessment of Mitochondrial DNA Structuring and Nursery Use in the North Atlantic Right Whale (*Eubalaena glacialis*)." *Canadian Journal of Zoology* 77: 1217–1222.

Marshall, T. C., J. Slate, L. E. B. Kruuk, and J. M. Pemberton. 1998. "Statistical Confidence for Likelihood-Based Paternity Inference in Natural Populations." *Molecular Ecology* 7: 639–655.

National Marine Fisheries Service. 2016. "Endangered and Threatened Species; Critical Habitat for Endangered North Atlantic Right Whale." *Federal Register* 81: 4838–4874.

Noren, D. P., and J. A. Mocklin. 2011. "Review of Cetacean Biopsy Techniques: Factors Contributing to Successful Sample Collection and Physiological and Behavioral Impacts." *Marine Mammal Science* 28: 154–199.

North Atlantic Right Whale Consortium. 2023. *North Atlantic Right Whale Consortium Identification Database 15-Feb-2023*. Anderson Cabot Center for Ocean Life at the New England Aquarium.

Pace, R. M., III, P. J. Corkeron, and S. D. Kraus. 2017. "State-Space Mark-Recapture Estimates Reveal a Recent Decline in Abundance of North Atlantic Right Whales." *Ecology and Evolution* 7: 8730–8741.

Payne, R., O. Brazier, E. M. Dorsey, J. S. Perkins, V. J. Rowntree, and A. Titus. 1983. "External Features in Southern Right Whales (*Eubalaena australis*) and Their Use in Identifying Individuals." In *Communication and Behavior of Whales*, edited by R. Payne, 371–445. American Association for the Advancement of Science.

Pearse, D. E., C. M. Eckerman, F. J. Janzen, and J. C. Avise. 2001. "A Genetic Analogue of 'Mark-Recapture' Methods for Estimating Population Size: An Approach Based on Molecular Parentage Assessments." *Molecular Ecology* 10: 2711–2718.

Pettis, H. M., R. M. Pace III, and P. K. Hamilton. 2023. "North Atlantic Right Whale Consortium 2022 Annual Report Card." Report to the North Atlantic Right Whale Consortium.

Posit Team. 2024. *RStudio: Integrated Development Environment for R*. Posit Software, PBC.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Reed, J., L. New, P. Corkeron, and R. Harcourt. 2022. "Multi-Event Modeling of True Reproductive States of Individual Female Right Whales Provides New Insights Into Their Decline." *Frontiers in Marine Science* 9: 994481.

Reeves, R. R., T. D. Smith, and E. A. Josephson. 2007. "Near-Annihilation of a Species: Right Whaling in the North Atlantic." In *The Urban Whale: North Atlantic Right Whales at the Crossroads*, edited by S. D. Kraus and R. M. Rolland, 39–74. Harvard University Press.

Ruzzante, D. E., G. R. McCracken, B. Førland, et al. 2019. "Validation of Close-Kin Mark–Recapture (CKMR) Methods for Estimating Population Abundance." *Methods in Ecology and Evolution* 10: 1445–1453.

Shaw, C. N., P. J. Wilson, and B. N. White. 2003. "A Reliable Molecular Method of Gender Determination for Mammals." *Journal of Mammalogy* 84: 123–128.

Silva, M. A., L. Steiner, I. Cascão, et al. 2012. "Winter Sighting of a Known Western North Atlantic Right Whale in the Azores." *Journal of Cetacean Research and Management* 12: 65–69.

Skaug, J. H. 2001. "Allele-Sharing Methods for Estimation of Population Size." *Biometrics* 57: 750–756.

Stan Development Team. 2024. "Stan Reference Manual, Version 2.36." https://mc-stan.org.

Walling, A., J. M. Pemberton, J. D. Hadfield, and L. E. B. Kruuk. 2010. "Comparing Parentage Inference Software: Reanalysis of a Red Deer Pedigree." *Molecular Ecology* 19: 1914–1928.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Number of genotyped mother-calf pairs available for paternity analysis by year. **Figure S2:** Number of paternities assigned by year using *MasterBayes* and a 95% criterion. **Figure S3:** The percentage of individuals used in paternity analyses that are genotyped at each number of loci.