Monitoring Wolf Packs by Counting Litters:

Sibling Reconstruction from Genetics to Genomics

A Thesis
Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science
with a
Major in Bioinformatics and Computational Biology
in the
College of Graduate Studies
University of Idaho
by

Heather R. Clendenin

Major Professors: Lisette Waits, Ph.D., Paul Hohenlohe, Ph.D.
Committee Members: David Ausband, Ph.D., Craig Miller, Ph.D.
Department Administrator: David Tank, Ph.D.

May 2019

Authorization to Submit Thesis

This thesis of Heather Clendenin, submitted for the degree of Master of Science with a Major in Bioinformatics and Computational Biology and titled "Monitoring Wolf Packs by Counting Litters: Sibling Reconstruction from Genetics to Genomics," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professors: _____ Date: _____
Lisette Waits, Ph.D.

_____ Date: _____
Paul Hohenlohe, Ph.D.

Committee Members: _____ Date: _____
David Ausband, Ph.D.

_____ Date: _____
Craig Miller, Ph.D.

Department
Administrator: _____ Date: _____
David Tank, Ph.D.

Abstract

Parameters of reproductive success are critical to the management of wildlife populations. Genetic monitoring can provide this demographic information when traditional methods aren't tractable. This study demonstrates a novel application of genetic data to estimate a minimum annual count of breeding packs of gray wolves (*Canis lupus*). Using tissue samples from wolves harvested in Idaho, 98 young of the year from 2015 and 205 from 2016 were genotyped at 18 microsatellite loci. Sibling groups for each cohort were reconstructed using COLONY, with full-sibling litters corresponding to unique packs. To assess the accuracy of relationship assignments, young of the year of known relationship from long-term study packs were added to the dataset (61 individuals from 2015 and 45 from 2016). Varied input parameters were used to evaluate the power of relationship assignments under real-world data constraints, providing insight into the use of sibship reconstruction as a tool to meet monitoring goals. Though all known relationship were correctly identified under these conditions, the number and size of assigned subgroups varied. Notably, the ability to discern between closely related non-siblings was diminished when the number of loci was reduced from 18 to 10. To further explore the impacts of different genotype data, we used RADseq to identify thousands of single nucleotide polymorphism (SNP) loci and repeated sibship analyses for 50 gray wolf YOY from 2014. Our results indicated that sibship analyses using SNP loci may be limited by missing data caused by DNA quality and quantity, and that strict filtering may yield inconsistent results. SNP and microsatellite datasets were generally concordant but produced some discrepancies in sibship assignments. To compare SNP-based sibship estimates against known relationships, we have also generated RADseq-derived SNPs in an analogous group of 86 red wolf (*Canis rufus*) YOY of known pedigree. Though the use of SNPs is increasing in genetic monitoring of wildlife, the strengths and trade-offs of these approaches—especially when working with legacy data sets or lower-quality DNA—should be considered in transitions to next generation methods.

Acknowledgements

My major advisors and committee members have been invaluable in the development of this project. I would like to express gratitude to Dr. David Ausband and Dr. Craig Miller for the enthusiasm, creativity, and feedback that helped bring this work to form, and to Dr. Paul Hohenlohe and Dr. Lisette Waits for providing me with the testing grounds to hone new skills and the opportunity to explore new challenges.

This study has been supported by the Laboratory for Ecological, Evolutionary and Conservation Genetics and funded by the Idaho Department of Fish and Game. I am grateful for the efforts of Idaho Department of Fish and Game field technicians and biologists for the collection of samples, and specifically Jennifer Struthers for collection, management and coordination of samples. I would like to express special appreciation to Kody Cochrell, Digpal Gour, Michelle Keyes, and Amanda Stahlke for assistance with genotyping and analyses, Dr. Jennifer Adams and Dr. Kimberly Andrews for sharing their technical skills and knowledge, and Jim Hayden for providing insight and guidance on this project.

Table of Contents

List of Figures

List of Tables

Statement of Contribution

Co-authors for these works are Jennifer Adams, David Ausband, Jim Hayden, Paul Hohenlohe, and Lisette Waits. All co-authors contributed suggestions and feedback on study design and participated in reviews of drafts and provided recommendations for improvements. Jennifer Adams conducted a pilot study of preliminary analyses with the microsatellite data from the 2014 cohort and provided guidance on use of long-term genetic monitoring data for both study species. David Ausband conducted parentage analyses for verification of results in chapter one and provided access to metadata. Jim Hayden was responsible for the original concept behind this study. Paul Hohenlohe provided feedback and guidance on the processing of sequence data and assisted in initial sequence data processing for gray wolves. Lisette Waits provided additional assistance with conceptual structure, development of methodologies, and acquired funding. Study design, analyses, figures, writing, and revisions were the work of the primary author, Heather Clendenin.

Chapter 1: Counting Litters Using Genetic Sibling Reconstruction of Harvested Wolves

Submitted to the Journal of Wildlife Management

Heather Clendenin,[1,2] Jennifer Adams,[1] David Ausband,[3] Jim Hayden,[3] Paul Hohenlohe, [2] and Lisette Waits[1]

Departments of [1]Fish and Wildlife Sciences and [2]Biological Sciences: Bioinformatics and Computational Biology, University of Idaho; [3]Idaho Department of Fish and Game

INTRODUCTION

Monitoring demographic parameters of populations is a crucial and demanding component of adaptive management. Comprehensive and cost-effective monitoring using traditional methods, such as direct observation and radio telemetry, can be intractable for species that are remotely located, behaviorally elusive, cryptic, or distributed across large geographic ranges (Ausband et al., 2014). Under these circumstances, genetic monitoring can be an efficient approach to obtaining reliable demographic information (De Barba et al. 2010, Stansbury et al. 2014). Genetic data can help identify species and individuals, provide estimates of population parameters, and offer insights into space use and connectivity (Schwartz et al. 2007, Paetkau et al. 2009, De Barba et al. 2010, Mumma et al. 2015, Micheletti and Storfer 2017). Measures of relatedness and genetic diversity can be used to reconstruct pedigrees, gain greater understanding of mating systems, assess population viability, and track quantitative traits (Thomas and Hill 2000, DeWoody 2005, Lucia and Keane 2011, Putnam and Ivy 2014, Gooley et al. 2017).

Relatedness analyses and kinship assignments compare genetic data from groups of individuals and use assumptions of identity by descent and Mendelian inheritance to assign relatedness within the context of relationship categories or as a continuous measure of

genetic similarity (DeWoody 2005). Assignment to family groups and other relatedness analyses have been applied to fisheries management, captive breeding programs, reintroductions, and translocations (Caballero et al. 2012, Fienieg and Galbusera 2013, Ottewell et al. 2014, Hopken et al. 2015, Caballero et al. 2017). These measures can also be used to track reproduction, estimate census and effective population sizes, and allow for close-kin genetic mark-recapture (Tokarska et al. 2009, Hauser et al. 2011, Fabbri et al. 2012, Artiles et al. 2015, Yu et al. 2015).

Sibship reconstruction is a relatedness assignment method that attempts to identify individuals of the same cohort that belong to common sibling groups (Almudevar and Anderson 2012). Sampling young of the year and identifying putative siblings has been used to identify breeders, trace the origins of individuals post-dispersal, and provide estimates of effective number of breeders or adult effective population size (Wang, 2005, Linløkken et al. 2016, Ackerman et al. 2017). Sibship reconstruction has also been valuable for tracking movement patterns and evaluating impacts of potential barriers (Kokuvo et al. 2007, Danancher et al. 2008, Whiteley et al. 2014, Quintela et al. 2016, Pope and Jha 2017). Estimates of effective population size based on the frequency of siblings identified within a sample have provided new insights into local extinction risk, sensitivity to environmental conditions, and resilience to change (Wang 2009, Straus et al. 2015, Ackerman et al. 2017, Whiteley et al. 2017, Bacles et al. 2018).

Reliable sibling reconstruction is the crux to meaningfully applying these methods in management and monitoring programs (Ackerman et al. 2017). Likelihood methods are commonly applied in the inference of sibships without parental information (Painter 1997, Wang 2004, Ashley et al. 2009). The certainty and reliability of these methods are

influenced by the number of genetic markers used and their variability within a population (DeWoody 2005). Using insufficiently informative markers can reduce the power of discernment and result in incorrect assignments, while failing to account for population allele frequencies can bias inferences through the disproportionate representation of alleles present in sampled families (Wang 2012). Adequate data collection is fundamental to accurate and reliable sibship assignments; however, resources available for monitoring and management of wild populations are often limited and careful consideration must be given to their allocation.

The management of gray wolves (*Canis lupus)* within the state of Idaho demonstrates such efforts to maximize the return on resources invested in monitoring. Gray wolves are managed as a game species in Idaho and are also monitored to ensure the population remains above recovery levels (Ausband et al. 2014, Stansbury et al. 2014, IDFG 2016). Responses to harvest within the conterminous U.S. vary from those documented in populations further north, making ongoing observation vital to responsive management actions (Ausband 2016). Though federal funding to support wolf monitoring has declined following delisting in the state, less intensive monitoring using carefully selected, complementary methods can be adequate for reporting a ground-level, minimum estimate of wolves given their abundance within the state (Ausband et al. 2014, Stansbury et al. 2014, Stenglein et al. 2011).

Existing studies have addressed questions relevant to wolf population ecology through direct observation, telemetry, hunter surveys, non-invasive genetic monitoring, and models of occupancy and vital rates (Rich et al. 2013, Bassing et al. 2015, Hindrikson et al. 2016, Stansbury et al. 2016, Granroth-Wilding et al. 2017). These studies have found that

the distribution and demography of wolves in this region can be driven by habitat quality and prey biomass, interspecific competition, and harvest pressure (Gude et al. 2012, Rich et al. 2012, Kittle et al. 2015, Bassing et al. 2016). These factors can also alter intra-group behavior and composition, which affect hunting success, territorial defense, mating strategies, survival of young, and total recruitment (MacNulty et al. 2009, Ruprecht et al. 2012, Ausband et al. 2015, Cassidy et al. 2016, Ausband 2018). Relatedness analyses conducted within this population have allowed not just for differentiation between degrees of relationships, but have permitted identification of finer relationship categories among individuals with the same coefficient of relatedness. This additional degree of refinement can offer insights into social structure and dynamics; siblings from separate reproductive events can be identified across years, patterns in inbreeding behavior can be revealed through knowledge of specific relationship categories, and extra-pair reproduction can be detected by relaxing assumptions of monogamy (Ausband 2018).

Taking advantage of harvest reporting and ongoing genetic monitoring, we propose use of sibling reconstruction as a novel method for estimating a minimum count of reproductively active wolf packs, while simultaneously generating an index of harvest vulnerability for young of the year at the pack level. Though sibship assignment requires more markers than individual identification, few additional resources are needed to add these analyses to the existing genetic monitoring program. Tissue samples and premolars are collected by Idaho Fish and Game personnel during harvest reporting, which facilitates genotyping and aging of harvested individuals. The number of litters affected by harvest can be found by reconstructing sibships among these harvested young of the year, which can be treated as a proxy for a minimum count of reproductively successful packs. Additionally,

reconstructing litters reveals the number of young harvested from individual packs, offering a picture of the distribution of harvest vulnerability.

In this project, we sought to assess the feasibility of using sibling reconstruction of harvested young of the year as a method for estimating a minimum count of reproductively successful wolf packs within the state of Idaho. To this end, individuals of known relationship (i.e. both siblings and non-siblings) from long-term study packs were embedded within a dataset of harvested young of the year to assess rates of correct assignment, and a series of parameter tests were run to assess the consistency of assignments. In the interest of extending these methods to other systems where monitoring resources may differ, we ran sibling reconstructions with various permutations of marker numbers and background information. We hypothesized that the most accurately and reliably reconstructed sibling configurations would come from runs using the longest run times and highest likelihood precision, and from the treatment using both the full marker set and background allele frequency data, and that the number of markers would have a larger bearing on the robustness of assignments than the use of background population allele frequency data.

STUDY AREA

The state of Idaho (216,632 km^2) contains many different landscapes, including mountainous forests, desert shrub, prairies, and open valleys. Elevations in the state range from 217m to more than 3,859m (United States Geological Survey, 2001). Public forests and private timber holdings, dominated by western red cedar (*Thuja plicata*) and western hemlock (*Tsuga heterophylla*), comprised most areas in northern Idaho. Management zones in central Idaho contained a mixture of wilderness areas, native prairies, and private

agricultural land, while areas in southern Idaho were predominantly private agricultural land. Annual precipitation ranged from <20 cm to >250cm, with temperatures ranging from -34°C to 38°C (Western Regional Climate Center, 2018). Gray wolves occurred and were subject to harvest throughout the state, with higher abundances in the northern and central portions. Hunting and trapping were regulated across 13 wolf management zones, subdivided into 98 game management units (Figure 1.1).

METHODS

Idaho Department of Fish and Game personnel collected tissue samples from voucher specimens provided during harvest reporting in 12 of Idaho's 13 wolf management zones (WMZs) during 2014 and 2015. Hunters and trappers are required to provide the hide and skull of harvested wolves during reporting and identify game management units (GMUs) hunted and date of harvest. Using premolars extracted during reporting, dentition analysis identified 94 harvested wolves as young of the year (YOY) in 2014, and 101 from 2015 (Figure 1.1).

In addition to harvested individuals, YOY of known relationship were sampled from long-term study packs within three of the management zones within the state (27 from 2014; 61 from 2015; Figure 1.1). Individuals within these packs have been monitored through annual field surveys and fecal DNA sampling at rendezvous sites, allowing for individual identification, pack assignment, and pedigree reconstruction (Stenglein et al. 2010, Ausband et al. 2014, Stansbury et al. 2016). The inclusion of YOY from these packs provided *a priori* knowledge of relationship categories, i.e. individuals within the same pack were known to be siblings, while individuals in different packs were known to be non-siblings. In

subsequently reconstructed sibgroups, these known relationships allowed us to identify incorrectly included or excluded individuals within this subset.

DNA was extracted from 20mg samples of tissue using Qiagen DNeasy Blood and Tissue kits, with negative controls included to test for contamination (Qiagen, Inc.). Eighteen dye-labelled nuclear DNA microsatellite loci were combined into two polymerase chain reaction (PCR) multiplexes with a product size of <300 bp (AHT103, AHT109, AHT121, AHT200, C05.377, C09.173, C37.172, Cxx.119, Cxx.250, FH2001, FH2004, FH2010, FH2054, FH2088, FH2137, FH2611, FH3725, Breen et al. 2001, Guyon et al. 2003, Holmes et al. 1994, Ostrander et al. 2017, Salim et al. 2007). Each multiplex PCR was run with a negative control to test for possible contamination of reagents. The 7-uL PCR reaction for multiplex 1 contained 0.09 uM 2670, 0.10 uM FH2611, 0.06 uM FH2088 and FH2054, 0.11 uM FH3725, 0.04 uM FH2137, 0.09 uM FH2001, 0.23 uM Cxx.119, 0.06 uM C09.173, 0.16 uM FH2004, and 3.5-uL of 1.5x concentrated Qiagen Master Mix, 0.7 -uL of 0.5x concentrated Qiagen Q Solution, and 2-uL DNA extract. The 7-uL PCR reaction for multiplex 2 contained 0.10 uM AHT121, 0.07 uM C37.172, 0.36 uM AHT103, 0.06 uM C05.377, 0.13 uM AHT109, 0.16 uM AHT200, 0.11 uM Cxx.250, 0.06 uM FH2010, 3.5-uL of 1x concentrated Qiagen Master Mix, 0.7-uL of 0.5x concentrated Qiagen Q Solution, and 2-uL DNA extract. The PCR profile for multiplex 1 had an initial 15 min denaturation step of 94°C, a touchdown of 13 cycles of 30 s at 94°C, 90 s at 62°C degrees with a decrease in annealing temperature of 0.4°C in each cycle, 1 min at 72°C degrees, followed by 17 cycles of 30 s at 94°C, 90 s at 57°C, 1 min at 72°C degrees. The profile of multiplex 2 varied in the annealing temperature, which began at 63°C degrees over 13 touchdown cycles and was 55°C in the subsequent 17 cycles.

PCR products were separated using an Applied Biosystems 3130xl capillary machine (Applied Biosystems Inc., Foster City, CA, USA), and genotypes were scored with GENEMAPPER 5.0 (Applied Biosystems Inc.). Samples were run in duplicate and repeated a third time when necessary to resolve genotype inconsistencies. Samples with consensus genotypes at 90% or more of the loci were used in sibling reconstruction.

Full-sibling relationships for these individuals were reconstructed using the software COLONY 2 (Jones and Wang, 2010). To assess sensitivity of Colony sibship assignments to input and parameter settings, we created three treatments (Table 1.1). Treatment 18BD (i.e. 18 microsatellite loci with background data) used genotype data from 18 microsatellites and allele frequencies and error rates based on tissue samples from 865 wolves collected in Idaho, genotyped between September 2013 and September 2017. Treatment 18NBD (i.e. 18 microsatellite loci with no background data) was also based on 18 genotyped loci, but allele frequencies and error rates were inferred only from individuals sampled within a given cohort. Treatment 10BD (i.e. 10 microsatellite loci with background data) used a subset of 10 loci and incorporated background population data into allele frequency and error rate estimates.

This full marker set was developed as two multiplexes for assigning pack membership, reconstructing familial relationships, and identifying dispersers among individuals detected during non-invasive surveying (Stenglein et al. 2011, Ausband et al. 2014, Stansbury et al, 2016). The first multiplex includes the 10 loci used in treatment 10BD, which has been the standard for individual ID of noninvasively monitored wolves in this state, while the addition of the second multiplex brings the total to 18 loci and allows improved resolution for assignment of parentage and identification of family groups. Allelic

richness, expected heterozygosity, and observed heterozygosity (Table S1.1, available online in Supporting Information).

Due to the nature of the optimization algorithm employed by Colony, reconstructions based on insufficiently informative markers may not necessarily converge on the same configuration of full-sibling families. Inconsistent and inaccurate assignments are more likely to occur when run length is shortened—either through adjusting the run-length parameter or by reducing the likelihood precision. As such, tracking changes in configuration likelihood and other measures of support while implementing a series of runs, including both replicates and a range of parameter settings, was used to reveal differences in power among treatments.

Each treatment was run five times to ascertain consistency of sibling assignments. Three sets of runs were executed with high full-likelihood precision, and differed only in run length (i.e. short, medium, and long). The remaining two replicated the short run with a different random number seed (e.g. 1234 in the original run, 12 in the replicate) and the other set at low full-likelihood precision. Parameter settings consistent across all runs included mating system (male and female monogamy; outbreeding), species (dioecious; diploid), analysis type (full-likelihood), and sibship scaling (not applied). Correct identification of known relationships, number of sibgroups, sizes of sibgroups, and assignments within sibgroups were evaluated for differences between runs and treatments.

Probabilities of inclusion (i.e. the probability that a full-sibling family contains only true siblings) and exclusion (i.e. the probability that the full-sibling family has not incorrectly excluded any true siblings) for sibgroups across runs, treatment types, and sizes

of sibgroups were evaluated using analysis of variance (ANOVA) and Tukey HSD tests for statistically significant differences using the program R, version 3.4.0 (R Core Team, 2017).

As an additional assessment of the plausibility of sibling assignments, the game management units (GMUs) and wolf management zones (WMZs) associated with the reporting of each harvested individual were compared for each member of a putative sibgroup. Management delineations do not necessarily correspond to territories or home ranges; however, geographical clustering can be expected of true siblings and disparate geographical locations of putative siblings may signal inaccurate assignments or incorrect reporting. While wolves have been described as dispersal pumps and reported territories vary in size between 33 km^2 to 4,335 km^2, typical dispersal distances average around 96.3 km (Boyd and Pletscher 1999, Jimenez et al. 2017). Though some exceptional distances have been recorded, both pre-dispersal forays and long-term dispersal typically occur at no earlier than 11 months of age, making long-distance travel unlikely within this age class (Boyd and Pletscher 1999, Mech and Boitani 2003, Vilà et al. 2003, Jimenez et al. 2017). Sibgroups of two or more members were categorized by those with all members associated with the same GMU, all members associated with the same WMZ but not within the same GMU, all members within adjacent WMZs, or members distributed in some other manner (Figure 1.1). Counts and percentages of non-singleton sibgroups within each category were generated, and categories were assessed and averaged for each run across treatments with rearranged sibling groupings.

Consistency of assignments were compared between treatments by identifying changes in the pair-by-pair assignment of individuals. For each cohort, all possible sibling pairings were scored per run as either siblings (1) or non-siblings (0), with these scores

summed across all runs. Pairwise assignments that remained consistent across all runs had summed scores of either 5 (always assigned as siblings) or 0 (never assigned as siblings), with intermediate scores indicating inconsistent assignments across runs.

RESULTS

Accurate Assignment of Known Relationships

Across all runs and treatments, all individuals of known relationship from the long-term study packs were correctly categorized, falling within six known sibling groups in the 2014 cohort and thirteen known sibling groups in 2015. No individuals known to be from separate packs were incorrectly assigned as siblings; however, two harvested individuals from the 2015 cohort that had not been reported among the long-term study packs were assigned to sibgroups with known-relationship wolves across all runs, with an additional harvested individual included under treatment 10BD. The full 18 microsatellite genotypes of these harvested individuals were compatible with their putative siblings (i.e. all alleles had previously been observed within this sibgroup), and subsequent parentage analyses confirmed shared parentage with assigned littermates for the two harvested individuals included among known sibgroups across all treatments (Table S1.2, available online in Supporting Information). However, parentage analysis indicated that, while the harvested individual added to a known sibgroup under treatment 10BD could have shared maternity with its putative siblings, it appeared to have been sired by an unrelated male (D. Ausband, IDFG, unpublished data). While the parentage analysis was not an independent evaluation of pack membership, it did provide additional support suggesting at least two harvested

individuals had true membership in long-term study packs despite lack of previous detection.

Number of Litters by Treatment

Among the 2014 cohort, there were no changes in the number of assigned sibling groups across runs in treatments 18BD and 18NBD (52 and 53 respectively). However, the total number of litters (46 and 47) were not consistent across runs in treatment 10BD with only 10 loci. In the 2015 cohort, all runs of treatments 18BD and 10BD generated the same number of litters (63 and 55 litters, respectively), while all but one run in 18NBD were the same (runs 1, 2, 3, and 5 all had 64 sibling groups, while run 4 had 63) (Figure 1.2).

For both cohorts, the maximum-likelihood configurations contained a smaller number of sibling groups in the treatment using 10 loci. Several individuals designated as singletons in treatments 18BD and 18NBD in the 2014 cohort were joined into pairs in treatment 10BD, while some singletons and pairs assigned in treatments 18BD and 18NBD in the 2015 cohort were assigned to larger groups in treatment 10BD (Figure 1.2, Table 1.2).

Spatial Distribution of Sibling Groups

Treatments 18BD and 18NBD performed similarly with respect to the spatial distribution of individuals assigned as siblings, with most putative sibling groups in both cohorts consisting of members all within the same game management unit or within the same wolf management zone. Treatment 10BD had the smallest fraction of sibgroups with members detected in the same management unit across cohorts, and showed the greatest disparity in spatial distribution of putative siblings (Figure 1.3).

Average Probability of Inclusion and Exclusion by Treatment

ANOVA results indicated that the only treatment with significantly different mean probabilities of inclusion and exclusion was 10BD (Table 1.3). Both the mean inclusion and exclusion probabilities for treatment 10BD were significantly lower than the other treatments in the 2014 cohort; in the 2015 cohort, treatment 10BD's mean probability of inclusion and exclusion were also lower than the other treatments, though only the probability of exclusion was significantly lower. No significant differences were detected between 18BD and 18NBD, nor among runs within treatments, in either cohort (Figure S1.1 & S1.2, available online in Supporting Information).

Consistency of Sibling Assignments

For the 2014 cohort, of all possible sibling pairings, there were 126 total sibling pairs observed in treatment 18BD, which all remained consistent across the five runs. For treatment 18NBD, 131 total pairs were observed in this cohort, with nine changes in pairings observed across runs (6 pairs were observed in only one of five runs, while 3 pairs observed in four of five total runs). 142 total pairs were observed in treatment 10BD with 7 rearranged pairs (4 pairs observed twice, three pairs observed three times). In the 2015 cohort, 205 total pairs with no rearrangements were observed in treatment 18BD, 210 total pairs were observed in treatment 18NBD with 13 pairs rearranged across runs (1 pair observed once, 6 pairs observed twice, and 6 pairs observed three times), and 222 pairs with no rearrangements were observed across all runs in treatment 10BD (Table 1.4).

DISCUSSION

Our study yielded promising results regarding the reliability and accuracy of sibling assignments and identification of sibling groups from harvested gray wolves using Colony. To our knowledge, our work represents the first effort to inform the management of a terrestrial game species using this method, and demonstrates a valuable new use of samples collected through harvest. Genetic sibling reconstruction of harvested young of the year can augment existing monitoring methods by providing a minimum count of reproductive wolf packs and an index of harvest vulnerability of young across packs (e.g. harvest was a source of mortality for 1-2 pups in most packs, but as many as 5 YOY were harvested from others). The value of estimating counts of parents and family groups through sibship reconstruction has been demonstrated in other systems, such as monitoring abundance of social bee colonies by identifying sisters among foraging workers and estimating the number of female sea turtles laying multiple clutches per season at the same nesting site (Toquenaga and Kokuvo 2010, Frey et al. 2013, Frey et al. 2014, Geib et al. 2015). Fisheries and aquaculture systems have used similar methods to shed light on genetic variability between age stages, assess stocking strategies, and reconstruct putative parental genotypes among externally fertilized species (Liu and Ely 2009, Li et al. 2013, Meraner et al. 2013, Hasanat et al. 2014). These applications, like ours, expand the information gained through genetic monitoring approaches.

Integrating multiple sources of information on wildlife populations can help to circumvent some of the challenges and weaknesses of individual monitoring methods and better capture demographic trends and responses to management actions (Ausband et al. 2014, Horne et al. 2018). In our system, with an existing genetic monitoring program,

sibling reconstruction requires few additional resources to estimate a minimum count of breeding packs within Idaho. The recovery status of wolves in the state was contingent upon maintaining a minimum number of breeding pairs, i.e. 10 pairs in either Idaho or Montana annually, or fewer than 15 pairs in either state for three consecutive years (U.S. Fish and Wildlife Service 2011). However, the 2009 U.S. Fish and Wildlife Service Wolf Delisting Rule defines breeding pairs as an adult male and an adult female wolf that have produced at least 2 pups that survived until December 31 of the year of their birth, during the previous breeding season (USFWS 2009). Given these stipulations, an estimate based on sibling reconstruction does not directly meet the former legal criteria regarding breeding pairs. However, this estimate could serve as a reasonable substitute for more resource-intensive methods of obtaining minimum counts of reproductive groups that have been used to validate probabilities of packs containing breeders and calibrate population models (Ausband et al. 2014, Mitchell et al. 2008, 2010, Gude et al. 2009). When assessed with other measures of population size, reproductive rates, and distribution of harvest pressure, this estimate can be treated as an index to track changes in the number of breeding pairs and population size.

In our study, we aimed to identify methods to optimize accuracy and certainty of sibling assignments. Sibship reconstruction accurately and consistently identified all *a priori* known relationships when using either 10 or 18 microsatellite loci, with or without background data on population-level allele frequencies. Systematic variation of input parameters allowed us to compare correct identification of known relationships and consistency of sibling reconstructions under conditions that may be encountered with other managed populations, allowing for assessment of parameter sensitivity and robustness of

assignments. We found that the number of markers used and the method of calculating allele frequencies had more impact on the accuracy and reliability of assignments than changes to the stringency of Colony parameter settings, such as run length and likelihood precision. Using a reduced set of 10 loci decreased the number of family groups detected and produced less credible individual assignments. When all 18 loci were used, sibling reconstructions identified a higher number of unique litters, with more probable assignments and greater spatial cohesion among putative siblings. Assignments using this full marker set were consistent across all runs when background population data was used to estimate allele frequencies, indicating further refinement and reliability of individual relationships under these conditions.

Though the true number of total sibling groups was not known for either cohort, identification of known relationships and other response variables can be used to guide considerations in the application of sibship reconstruction. Notably, our results suggested that there were inaccurate individual assignments when using the restricted set of 10 loci. Using insufficient marker data appears to have produced errors of false inclusion. Reconstructions in this treatment type had both significantly lower probabilities of inclusion and exclusion, and also generated significantly lower total group counts relative to those using all 18 loci. Across both cohorts, several singletons and putative sibling pairs identified using the full marker set were joined into larger groups in assignments using fewer markers. We believe these differences were caused by reduced resolution, resulting in loss of discernment between non-siblings— whether non-relatives with similar genotypes or relatives of other relatedness categories. Though the correct identification of known relationships showed that assignments were not arbitrary when using the reduced set of loci,

total assignments appeared to err low, with presumably erroneous assignments primarily occurring through over-joining—a trend observed elsewhere under conditions where, due to marker information and family structure, non-siblings have genotypes consistent with full siblings by chance (Chapman et al. 2003, Wang and Santure 2009, Lepais et al. 2010). This would be a probable explanation for an assignment unique to this treatment, in which a harvested individual was added to a litter of known-siblings, was not confirmed by subsequent parentage analyses (D. Ausband, unpublished data). Interpreting the reduced number of sibgroups as an error due to loss of resolution is further supported by the geographical distribution of putative siblings. In assignments based on 10 loci, a larger percent of sibling groups contained putative members with less spatial proximity to each other than siblings identified in other treatments (Figure 1.3).

In contrast to the reconstructions using 10 loci, the congruency in results across treatments using the full marker set lends assurance that the 18 microsatellites loci used and the threshold set for missing data provide enough information and power to make these assessments, with or without background allele frequency data. All reconstructions using 18 microsatellites resulted in similar total counts of sibling groups, comparable mean probabilities of inclusion and exclusion, and higher percentages of sibling groups with members reported within the same or adjacent management areas. Given these improvements, the increased total counts of sibling groups in these treatments likely represent unique reproductive packs, which reconstructions using fewer markers were not able to discriminate. Uncertainty in individual assignments was further reduced when available background data was used, producing consistent sibling assignments across runs. Overall, using both the complete marker set and applying existing genetic monitoring data

produced assignments in which both group counts, as well as individual relationships, were the most reliable.

It is important to note that our proposed monitoring method of using the number of reconstructed litters from harvested young of the year as a proxy for the number of reproductive packs only provides a minimum count and will be limited by harvest rates and distribution, and that variation in the quality of sample preservation can reduce successful genotyping rates and influence detection probabilities. Interestingly, our estimates of minimum pack counts were close to annual estimates of reproducing packs based on other survey methods, but lower than these counts for both years (Table 1.5).

Though harvest as a monitoring tool entails different challenges in assessing detection probabilities, there are several advantages to this source of data (Leclerc et al. 2016). Hunter surveys and harvest reporting provide valuable information on wildlife populations and engage stakeholders in the monitoring process (Rich et al. 2013, Leclerc et al. 2016). Additionally, genetic analyses based on harvested individuals can complement and contrast non-invasive genetic sampling. Non-invasive sampling is limited by agency time and resources, resulting in patchiness in spatial detection patterns. Though harvest can also exhibit spatial bias, these biases are not likely to be the same, allowing the geographical distribution of harvest-based sampling to supplement that of agency monitoring. This can improve the accuracy of individual assignment and aging relative to non-invasive sampling, as tissue generally allows for greater genotyping success and dentition analysis involves less subjectivity than approximating the age of a sampled individual via relative size of scat (Weaver and Fritts 1979, Gipson et al. 2000). Recognizing this valuable resource, sibling

reconstruction of harvested young of the year can take advantage of an opportunistic source of data to provide further information on managed populations.

MANAGEMENT IMPLICATIONS

We have demonstrated the accuracy and reliability of sibling reconstruction of harvested young of the year to estimate minimum annual counts of reproductive wolf packs and the harvest vulnerability of young across packs. Taken in conjunction with other monitored population parameters, these counts can be used to estimate a baseline rate of reproduction, detect the distribution of harvest vulnerability on specific packs, and offer an index of survival of young, which can subsequently be incorporated into traditional monitoring approaches to provide insights into population abundance, vital rates, and spatial distribution. Other ecological and evolutionary dynamics can also be explored by extending these analyses to estimates of population genetic parameters, such as effective population size and measures of gene flow, or by using sibling reconstruction to understand mating systems and family structure.

In exploring the reliability and robustness of sibling assignments when constrained by available background data and number of markers, we hope to encourage the application of this practice in other systems where resources and research questions may differ. In our system, we found our set of 18 microsatellite loci capable of producing relatively consistent total counts of sibling groups across the sampled cohorts. The reliability of these assignments was further improved by incorporating background population data to estimate allele frequencies, resulting in uniform individual assignments across all Colony runs. Though this study benefitted from an existing genetic monitoring program, selecting markers for new studies in the absence of legacy datasets and when few initial parameters

are known will require additional resources. Application to other systems will be affected not only by background population knowledge and marker number, but the allelic richness and heterozygosity of the markers used, and can be guided by theoretical models and available information on populations of the same species (Waits et al. 2001).

LITERATURE CITED

Ackerman, M. W., B. K. Hand, R. K. Waples, G. Luikart, R. S. Waples, C. A. Steele, B. A. Garner, J. McCane, and M. R. Campbell. 2017. Effective Number of Breeders from Sibship Reconstruction: Empirical Evaluations Using Hatchery Steelhead. Evolutionary Applications 10 (2): 146–60.

Almudevar, A., and E. C. Anderson. 2012. A New Version of PRT Software for Sibling Groups Reconstruction with Comments Regarding Several Issues in the Sibling Reconstruction Problem. Molecular Ecology Resources 12 (1): 164–78.

Artiles, A., R. Cobo, L. Benítez, L. Pérez, and G. Espinosa. 2015. Assessment of Genetic Variation and Productive Markers through Four Progenies of the First Introduced Stock of Cultured Shrimp *Penaeus (Litopenaeus) Vannamei* in Cuba. International Journal of Aquaculture. 5.

Ashley, M. V., T. Y. Berger-Wolf, I. C. Caballero, W. Chaovalitwongse, B. DasGupta, and S. I. Sheikh. 2008. Full Sibling Reconstruction in Wild Populations From Microsatellite Genetic Markers. Computational Biology: New Research. Nova Science, Hauppauge, 231-258.

Ausband, D. E. 2016. Gray Wolf Harvest in Idaho. Wildlife Society Bulletin 40 (3): 500–505.

Ausband, D. E. 2018. Multiple Breeding Individuals within Groups in a Social Carnivore. Journal of Mammalogy 99 (4): 836–844.

Ausband, D. E., M. S. Mitchell, C. R. Stansbury, J. L. Stenglein, and L. P. Waits. 2017. Harvest and Group Effects on Pup Survival in a Cooperative Breeder. Proceedings of the Royal Society B: Biological Sciences 284 (1855): 20170580.

Ausband, D. E., M. S. Mitchell, and L. P. Waits. 2017. Effects of Breeder Turnover and Harvest on Group Composition and Recruitment in a Social Carnivore. Journal of Animal Ecology 86 (5): 1094–1101.

Ausband, D. E., L. N. Rich, E. M. Glenn, M. S. Mitchell, P. Zager, D. A. W. Miller, L. P. Waits, B. B. Ackerman, and C. M. Mack. 2014. Monitoring Gray Wolf Populations Using Multiple Survey Methods. The Journal of Wildlife Management 78 (2): 335–346.

Ausband, D. E., C. R. Stansbury, J. L. Stenglein, J. L. Struthers, and L. P. Waits. 2015. Recruitment in a Social Carnivore before and after Harvest. Animal Conservation 18 (5): 415–423.

Bacles, C. F. E., C. Bouchard, F. Lange, A. Manicki, C. Tentelier, and O. Lepais. 2018. Estimating the Effective Number of Breeders from Single Parr Samples for Conservation Monitoring of Wild Populations of Atlantic Salmon Salmo Salar. Journal of Fish Biology 92 (3): 699–726.

Bassing, S. B., M. S. Mitchell, D. E. Ausband, L. P. Waits, and G. Hale. 2015. Developing a Monitoring Framework to Estimate Wolf Distribution and Abundance in Southwest Alberta. Intermountain Journal of Sciences 21 (1–4): 56–57.

Bassing, S. B., M. S. Mitchell, P. M. Lukacs, D. E. Ausband, L. P. Waits, and G. Hale. 2016. Wolf Pack Distribution in Relation to Heavy Harvest in Southwest Alberta. Intermountain Journal of Sciences 22 (4): 96.

Boyd, D. K., and D. H. Pletscher. 1999. Characteristics of Dispersal in a Colonizing Wolf Population in the Central Rocky Mountains. The Journal of Wildlife Management 63 (4): 1094–1108.

Breen, M., S. Jouquand, C. Renier, C. S. Mellersh, C. Hitte, N. G. Holmes, A. Chéron, N. Suter, F. Vignaux, A. E. Bristow, C. Priat, E. McCann, C. Andre´, S. Boundy, P. Gitsham, R. Thomas, W. L. Bridge, H. F. Spriggs, E. J. Ryder, A. Curson, J. Sampson, E. A. Ostrander, M. M. Binns, and F. Galibert. 2001. Chromosome-Specific Single-Locus FISH Probes Allow Anchorage of an 1800-Marker Integrated Radiation-Hybrid/Linkage Map of the Domestic Dog Genome to All Chromosomes. Genome Research 11 (10): 1784–1795.

Caballero, S., D. Cardenosa, G. Soler, and J. Hyde, 2012. Application of multiplex PCR approaches for shark molecular identification: feasibility and applications for fisheries management and conservation in the Eastern Tropical Pacific. Molecular Ecology Resources, 12(2): 233-237.

Caballero, A., I. Bravo, and J. Wang. 2017. Inbreeding Load and Purging: Implications for the Short-Term Survival and the Conservation Management of Small Populations. Heredity 118 (2): 177–185.

Cassidy, K. A., D. W. Smith, L. D. Mech, D. R. MacNulty, D. R. Stahler, and M. C. Metz. 2016. Territoriality and Inter-Pack Aggression in Gray Wolves: Shaping a Social Carnivore's Life History. Yellowstone Science 24 (1): 37-42.

Chapman, R. E., J. Wang, and A. F. Bourke, 2003. Genetic analysis of spatial foraging patterns and resource sharing in bumble bee pollinators. Molecular Ecology, 12(10): 2801-2808.

Danancher, D., J. I. Izquierdo, and E. Garcia-Vazquez. 2008. Microsatellite Analysis of Relatedness Structure in Young of the Year of the Endangered *Zingel Asper (Percidae)* and Implications for Conservation. Freshwater Biology 53 (3): 546–557.

De Barba, M., L. P. Waits, E. O. Garton, P. Genovesi, E. Randi, A. Mustoni, and C. Groff. 2010. The Power of Genetic Monitoring for Studying Demography, Ecology and Genetics of a Reintroduced Brown Bear Population. Molecular Ecology 19 (18): 3938–3951.

DeWoody, J. Andrew. 2005. Molecular Approaches to the Study of Parentage, Relatedness, and Fitness: Practical Applications for Wild Animals. The Journal of Wildlife Management 69 (4): 1400–1418.

Fabbri, E., R. Caniglia, N. Mucci, H. P. Thomsen, K. Krag, C. Pertoldi, V. Loeschcke, and E. Randi. 2012. Comparison of Single Nucleotide Polymorphisms and Microsatellites in Non-Invasive Genetic Monitoring of a Wolf Population. Archives of Biological Sciences 64 (1): 321–335.

Fienieg, E. S. and P. Galbusera, 2013. The use and integration of molecular DNA information in conservation breeding programmes: a review. Journal of Zoo and Aquarium Research, 1(2): 44-51.

Frey, A., P. H. Dutton, and G. H. Balazs. 2013. Insights on the Demography of Cryptic Nesting by Green Turtles *(Chelonia Mydas)* in the Main Hawaiian Islands from Genetic Relatedness Analysis. Journal of Experimental Marine Biology and Ecology 442: 80–87.

Frey, A., P. H. Dutton, D. J. Shaver, J. S. Walker, and C. Rubio. 2014. Kemp's Ridley *Lepidochelys Kempii* Nesting Abundance in Texas, USA: A Novel Approach Using Genetics to Improve Population Census. Endangered Species Research 23 (1): 63–71.

Geib, J. C., J. P. Strange, and C. Galen. 2015. Bumble Bee Nest Abundance, Foraging Distance, and Host-Plant Reproduction: Implications for Management and Conservation. Ecological Applications 25 (3): 768–778.

Gipson, P. S., W. B. Ballard, R. M. Nowak, and L. D. Mech. 2000. Accuracy and Precision of Estimating Age of Gray Wolves by Tooth Wear. The Journal of Wildlife Management 64 (3): 752–758.

Gooley, R., C. J. Hogg, K. Belov, and C. E. Grueber. 2017. No Evidence of Inbreeding Depression in a Tasmanian Devil Insurance Population despite Significant Variation in Inbreeding. Scientific Reports 7 (1): 1830.

Granroth-Wilding, H., C. Primmer, M. Lindqvist, J. Poutanen, O. Thalmann, J. Aspi, J. Harmoinen, I. Kojola, and T. Laaksonen. 2017. Non-Invasive Genetic Monitoring Involving Citizen Science Enables Reconstruction of Current Pack Dynamics in a Re-Establishing Wolf Population. BMC Ecology 17 (1): 44.

Gude, J. A., M. S. Mitchell, D. E. Ausband, C. A. Sime, and E. E. Bangs. 2009. Internal Validation of Predictive Logistic Regression Models for Decision-Making in Wildlife Management. Wildlife Biology 15 (4): 352–369.

Gude, J. A., M. S. Mitchell, R. E. Russell, C. A. Sime, E. E. Bangs, L. D. Mech, and R. R. Ream. 2012. Wolf Population Dynamics in the U.S. Northern Rocky Mountains Are Affected by Recruitment and Human-Caused Mortality. The Journal of Wildlife Management 76 (1): 108–18.

Guyon, R., T. D. Lorentzen, C. Hitte, L. Kim, E. Cadieu, H. G. Parker, P. Quignon, et al. 2003. A 1-Mb Resolution Radiation Hybrid Map of the Canine Genome. Proceedings of the National Academy of Sciences 100 (9): 5296–5301.

Hasanat, A., F. A. Mollah, and S. Alam. 2014. Assessment of Genetic Diversity in Wild and Hatchery Populations of *Mrigal Cirrhinus Cirrhosus (Hamilton-Buchanan)* Using Allozyme Markers. International Journal of Fisheries and Aquatic Studies 1 (4): 24–31.

Hauser, L., M. Baird, R. Hilborn, L. W. Seeb, and J. E. Seeb. 2011. An Empirical Comparison of SNPs and Microsatellites for Parentage and Kinship Assignment in a Wild Sockeye Salmon *(Oncorhynchus Nerka)* Population. Molecular Ecology Resources 11: 150-161

Hindrikson M., J. Remm, M. Pilot, R. Godinho, A. V. Stronen, L. Baltrūnaité, S. D. Czarnomska, J. A. Leonard, E. Randi, C. Nowak, and M. Åkesson, 2017. Wolf Population Genetics in Europe: A Systematic Review, Meta-Analysis and Suggestions for Conservation and Management. Biological Reviews 92 (3): 1601–1629.

Holmes, N. G., N. J. Strange, M. M. Binns, C. S. Mellersh, and J. Sampson. 1994. Three Polymorphic Canine Microsatellites. Animal Genetics 25 (3): 200–200.

Hopken, M., T. Lum, P. Meyers, and A. Piaggio. 2015. Molecular Assessment of Translocation and Management of an Endangered Subspecies of White-Tailed Deer *(Odocoileus Virginianus)*. Conservation Genetics 16 (3): 635–647.

Horne, J. S., D. E. Ausband, M. A. Hurley, J. Struthers, J. E. Berg, and K. Groth. 2019. Integrated Population Model to Improve Knowledge and Management of Idaho Wolves. The Journal of Wildlife Management 83 (1): 32-42.

Idaho Department of Fish and Game. 2016. Statewide report surveys and inventory: Wolf surveys and inventories, wolf population status and trend studies. Boise, Idaho, USA.

Jimenez, M. D., E. E. Bangs, D. K. Boyd, D. W. Smith, S. A. Becker, D. E. Ausband, S. P. Woodruff, E. H. Bradley, J. Holyan, and K. Laudon. 2017. Wolf Dispersal in the Rocky Mountains, Western United States: 1993–2008. The Journal of Wildlife Management, 81 (4): 581-592

Jones, O. R., and J. Wang. 2010. COLONY: A Program for Parentage and Sibship Inference from Multilocus Genotype Data. Molecular Ecology Resources 10 (3): 551–555.

Kittle, A. M., M. Anderson, T. Avgar, J. A. Baker, G. S. Brown, J. Hagens, E. Iwachewski, S. Moffatt, A. Mosser, B. R. Patterson, and D. E. Reid, 2015. Wolves Adapt Territory Size, Not Pack Size to Local Habitat Quality. Journal of Animal Ecology 84 (5): 1177–1186.

Kokuvo, N., Y. Toquenaga, and K. Goka. 2007. A Simple Visualization Method to Reconstruct Nest-Mate Patterns among Bumble Bees *(Hymenoptera: Apidae)* Using Genetic Data. Applied Entomology and Zoology 42 (1): 137–141.

Leclerc, M., J. Van de Walle, A. Zedrosser, J. E. Swenson, and F. Pelletier. 2016. Can Hunting Data Be Used to Estimate Unbiased Population Parameters? A Case Study on Brown Bears. Biology Letters 12 (6): 20160197.

Lepais, O., B. Darvill., S. O'Connor, J. L. Osborne, R. A. Sanderson, J. Cussans, L. Goffe, and D. Goulson. 2010. Estimation of bumblebee queen dispersal distances using sibship reconstruction method. Molecular Ecology, 19(4), pp.819-831.

Li, R., Q. Li, and C. Wang. 2013. Sibship Reconstruction and Effective Population Size Estimation in Mass Spawning Ark Shell, *Scapharca Broughtonii* Based on Microsatellite Analysis. Genes & Genomics 35 (6): 703–708.

Liu, G., A. B. Shafer, W. Zimmermann, D. Hu, W. Wang, H. Chu, J. Cao, and C. Zhao. 2014. Evaluating the Reintroduction Project of Przewalski's Horse in China Using Genetic and Pedigree Data. Biological Conservation 171: 288–298.

Liu, J. X., and B. Ely. 2009. Sibship Reconstruction Demonstrates the Extremely Low Effective Population Size of Striped Bass *Morone Saxatilis* in the Santee–Cooper System, South Carolina, USA. Molecular Ecology 18 (19): 4112–4120.

Lucia, K. E., and B. Keane. 2011. A Field Test of the Effects of Familiarity and Relatedness on Social Associations and Reproduction in Prairie Voles. Behavioral Ecology and Sociobiology 66 (1): 13–27.

MacNulty, D. R., D. W. Smith, J. A. Vucetich, L. D. Mech, D. R. Stahler, and C. Packer. 2009. Predatory Senescence in Ageing Wolves: Ageing Limits Wolf Predation. Ecology Letters 12 (12): 1347–1356.

Mech, L D., and L. Boitani. 2003. Wolves: Behavior, Ecology, and Conservation. Chicago and London: University of Chicago.

Meraner, A., G. Unfer, and A. Gandolfi. 2013. Good News for Conservation: Mitochondrial and Microsatellite DNA Data Detect Limited Genetic Signatures of Inter-Basin Fish Transfer in *Thymallus Thymallus (Salmonidae) from* the Upper Drava River. Knowledge and Management of Aquatic Ecosystems, 409: 01.

Micheletti, S. J., and A. Storfer. 2017. An Approach for Identifying Cryptic Barriers to Gene Flow That Limit Species' Geographic Ranges. Molecular Ecology 26 (2): 490–504.

Mitchell, M. S., J. A. Gude, D. E. Ausband, C. A. Sime, E. E. Bangs, M. D. Jimenez, C. M. Mack, T. J. Meier, M. S. Nadeau, and D. W. Smith. 2010. Temporal Validation of an Estimator for Successful Breeding Pairs of Wolves *Canis Lupus* in the U.S. Northern Rocky Mountains. Wildlife Biology 16 (1): 101–106.

Mitchell, M. S., D. E. Ausband, C. A. Sime, E. E. Bangs, J. A. Gude, J. A. Jimenez, C. M. Mack, T. J. Meier, M. S. Nadeau, and D. W. Smith. 2008. Estimation of successful breeding pairs for wolves in the Northern Rocky Mountains, USA. The Journal of Wildlife Management, 72(4): 881-891

Mumma, M. A., C. Zieminski, T. K. Fuller, S. P. Mahoney, and L. P. Waits. 2015. Evaluating Noninvasive Genetic Sampling Techniques to Estimate Large Carnivore Abundance. Molecular Ecology Resources 15 (5): 1133–1144

Ostrander, E. A., R. K. Wayne, A. H. Freedman, and B. W. Davis. 2017. Demographic History, Selection and Functional Diversity of the Canine Genome. Nature Reviews Genetics 18 (12): 705–720.

Ottewell, K., J. Dunlop, N. Thomas, K. Morris, D. Coates, and M. Byrne. 2014. Evaluating Success of Translocations in Maintaining Genetic Diversity in a Threatened Mammal. Biological Conservation 171: 209–219.

Paetkau, D., E. Vázquez-Domínguez, N. I. Tucker, and C. Moritz. 2009. Monitoring Movement into and through a Newly Planted Rainforest Corridor Using Genetic Analysis of Natal Origin. Ecological Management & Restoration 10 (3): 210–216.

Painter, I. 1997. Sibship Reconstruction without Parental Information. Journal of Agricultural, Biological, and Environmental Statistics 2 (2): 212–229.

Pope, N. S., and S. Jha. 2017. Inferring the Foraging Ranges of Social Bees from Sibling Genotypes Sampled across Discrete Locations. Conservation Genetics 18 (3): 645–658.

Putnam, A. S., and J. A. Ivy. 2014. Kinship-Based Management Strategies for Captive Breeding Programs When Pedigrees Are Unknown or Uncertain. The Journal of Heredity 105 (3): 303–311.

Quintela, M., V. Wennevik, A. G. Sørvik, Ø. Skaala, O. T. Skilbrei, K. Urdal, B. T. Barlaup, and K. A. Glover. 2016. Siblingship Tests Connect Two Seemingly Independent Farmed Atlantic Salmon Escape Events. Aquaculture Environment Interactions 8: 497–509.

Rich, L. N., M. S. Mitchell, J. A. Gude, and C. A. Sime. 2012. Anthropogenic Mortality, Intraspecific Competition, and Prey Availability Influence Territory Sizes of Wolves in Montana. Journal of Mammalogy 93 (3): 722–731.

Rich, Li. N., R. E. Russell, E. M. Glenn, M. S. Mitchell, J. A. Gude, K. M. Podruzny, C. A. Sime, K. Laudon, D. E. Ausband, and J. D. Nichols. 2013. Estimating Occupancy and Predicting Numbers of Gray Wolf Packs in Montana Using Hunter Surveys. The Journal of Wildlife Management 77 (6): 1280–1289.

Ruprecht, J. S., D. E. Ausband, M. S. Mitchell, E. O. Garton, and P. Zager. 2012. Homesite Attendance Based on Sex, Breeding Status, and Number of Helpers in Gray Wolf Packs. Journal of Mammalogy 93 (4): 1001–1005.

Salim, D. C., A. A. Akimoto, C. B. Carvalho, S. F. Oliveira, C. K. Grisolia, J. R. Moreira, and M. N. Klautau-Guimarães. 2007. Genetic Variability in Maned Wolf Based on Heterologous Short-Tandem Repeat Markers from Domestic Dog. Genetics and Molecular Research, 6(2): 348-357.

Schwartz, M. K., G. Luikart, and R. S. Waples. 2007. Genetic Monitoring as a Promising Tool for Conservation and Management. Trends in Ecology & Evolution 22 (1): 25–33.

Stansbury, C. R., D. E. Ausband, P. Zager, C. M. Mack, C. R. Miller, M. W. Pennell, and L. P. Waits. 2014. A Long-Term Population Monitoring Approach for a Wide-Ranging Carnivore: Noninvasive Genetic Sampling of Gray Wolf Rendezvous Sites in Idaho, USA. The Journal of Wildlife Management 78 (6): 1040–1049.

Stansbury, C. R., D. E. Ausband, P. Zager, C. M. Mack, and L. P. Waits. 2016. Identifying Gray Wolf Packs and Dispersers Using Noninvasive Genetic Samples. The Journal of Wildlife Management 80 (8): 1408-1419

Stenglein, J. L., L. P. Waits, D. E. Ausband, P. Zager, and C. M. Mack. 2010. Efficient, Noninvasive Genetic Sampling for Monitoring Reintroduced Wolves. The Journal of Wildlife Management 74 (5): 1050–1058.

Stenglein, J. L., L. P. Waits, D. E. Ausband, P. Zager, and C. M. Mack. 2011. Estimating Gray Wolf Pack Size and Family Relationships Using Noninvasive Genetic Sampling at Rendezvous Sites. Journal of Mammalogy 92 (4): 784–795.

Straus, K. M., B. Vadopalas, J. P. Davis, and C. S. Friedman. 2015. Reduced Genetic Variation and Decreased Effective Number of Breeders in Five Year-Classes of Cultured Geoducks *(Panopea Generosa).* Journal of Shellfish Research 34 (1): 163–169.

Thomas, S. C., and W. G. Hill. 2000. Estimating Quantitative Genetic Parameters Using Sibships Reconstructed from Marker Data. Genetics 155 (4): 1961–1972.

Tokarska, M., T. Marshall, R. Kowalczyk, J. M. Wójcik, C. Pertoldi, T. N. Kristensen, V. Loeschcke, V. R. Gregersen, and C. Bendixen. 2009. Effectiveness of Microsatellite and SNP Markers for Parentage and Identity Analysis in Species with Low Genetic Diversity: The Case of European Bison. Heredity 103 (4): 326–32.

Toquenaga, Y., and N. Kokuvo. 2010. Full-Sib Reconstruction in Haplodiploid Populations. Applied Entomology and Zoology 45 (1): 59–64.

U.S. Fish and Wildlife Service [USFWS]. 2009. Endangered and threatened wildlife and plants; final rule to identify the northern Rocky Mountain population of gray wolf as a distinct population segment and to revise the list of Endangered and Threatened wildlife. Federal Register 74:15123–15188.

U.S. Fish and Wildlife Service [USFWS]. 2011. Reissuance of Final Rule to Identify the Northern Rocky Mountain Population of Gray Wolf as a Distinct Population Segment and to Revise the List of Endangered and Threatened Wildlife. Federal Register 76:25590–25592.

Vilà, C., A. K. Sundqvist, Ø. Flagstad, J. Seddon, S. B. merfeldt, I. Kojola, A. Casulli, H. Sand, P. Wabakken, and H. Ellegren. 2003. Rescue of a Severely Bottlenecked Wolf (Canis Lupus) Population by a Single Immigrant. Proceedings. Biological Sciences 270 (1510): 91–97.

Waits, L. P., G. Luikart, and P. Taberlet. 2001. Estimating the Probability of Identity among Genotypes in Natural Populations: Cautions and Guidelines. Molecular Ecology 10: 249–256.

Wang, J. 2004. Sibship Reconstruction from Genetic Data with Typing Errors. Genetics 166 (4): 1963–1979.

Wang, J. 2005. Estimation of Effective Population Sizes from Data on Genetic Markers. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 360 (1459): 1395–1409.

Wang, J. 2009. A New Method for Estimating Effective Population Sizes from a Single Sample of Multilocus Genotypes. Molecular Ecology 18 (10): 2148–2164.

Wang, J. 2012. Computationally Efficient Sibship and Parentage Assignment from Multilocus Marker Data. Genetics 191 (1): 183–194.

Wang, J. and Santure, A. W., 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics, 181 (4): 1579-1594.

Weaver, J., and S. H. Fritts. 1979. Comparison of Coyote and Wolf Scat Diameters. Journal of Wildlife Management 43 (3): 3.

Whiteley, A. R., J. A. Coombs, B. H. Letcher, and K. H. Nislow. 2014. Simulation and Empirical Analysis of Novel Sibship-Based Genetic Determination of Fish Passage. Canadian Journal of Fisheries and Aquatic Sciences 71 (11): 1667–79.

Whiteley, A. R., J. A. Coombs, M. J. O'Donnell, K. H. Nislow, and B. H. Letcher. 2017. Keeping Things Local: Subpopulation Nb and Ne in a Stream Network with Partial Barriers to Fish Migration. Evolutionary Applications 10 (4): 348–65.

Yu, G. C., Q. Z. Tang, K. R. Long, T. D. Che, M. Z. Li, and S. R. Shuai. 2015. Effectiveness of Microsatellite and Single Nucleotide Polymorphism Markers for Parentage Analysis in European Domestic Pigs. Genetics and Molecular Research 14 (1): 1362–1370.

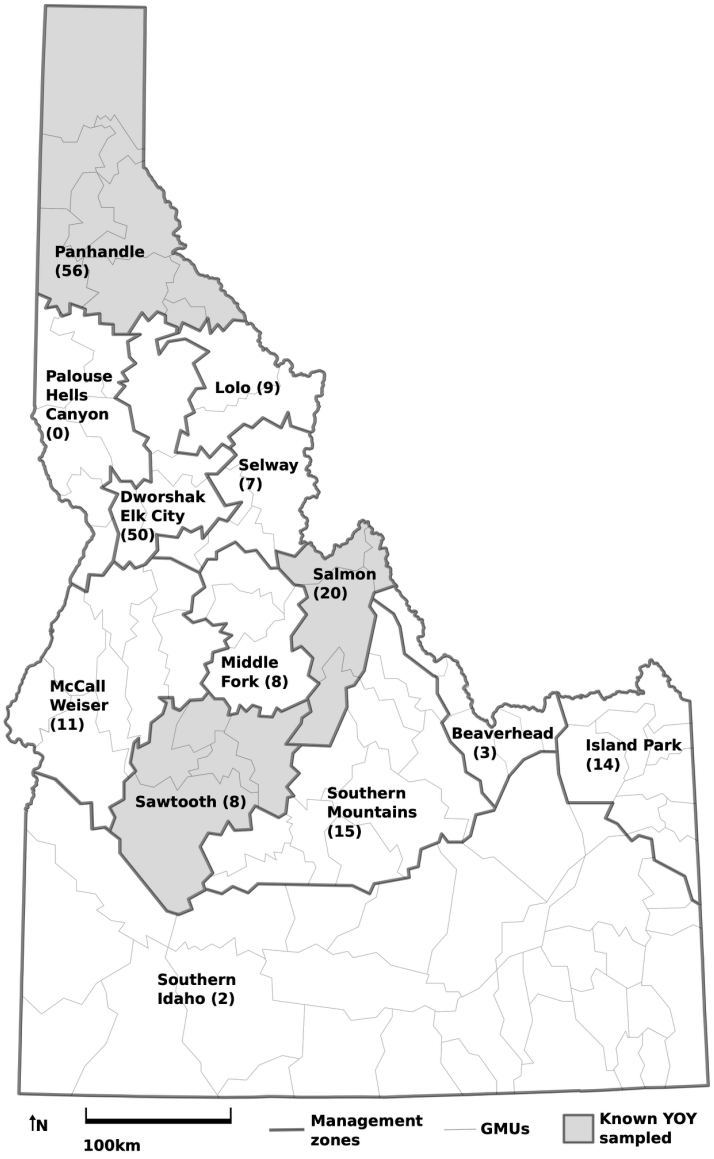

Figure 1.1: Spatial distribution of harvested gray wolf YOY from 2014 and 2015, as reported within the 13 wolf management zones in the state of Idaho. Number in parentheses following wolf management zone name indicates the total count of harvested YOY used in sibling reconstructions across both years. Individuals of known relationship category were sampled in the shaded management zones. These individuals were not included in the depicted counts, but included 25 YOY from the Panhandle wolf management zone, 34 from Salmon, and 32 from Sawtooth.

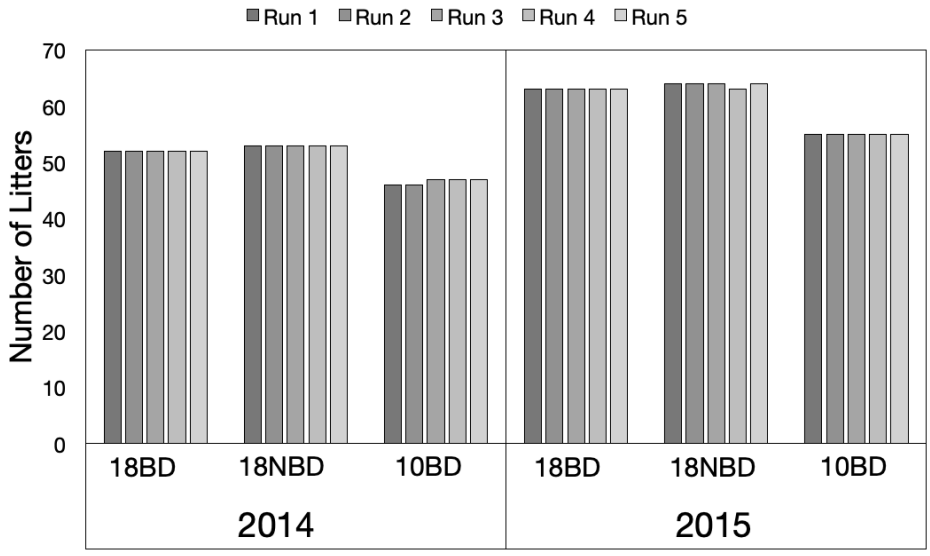

Figure 1.2: Counts of total gray wolf litters (i.e. putative sibling groups) assigned by cohort year, treatment type, and run number.

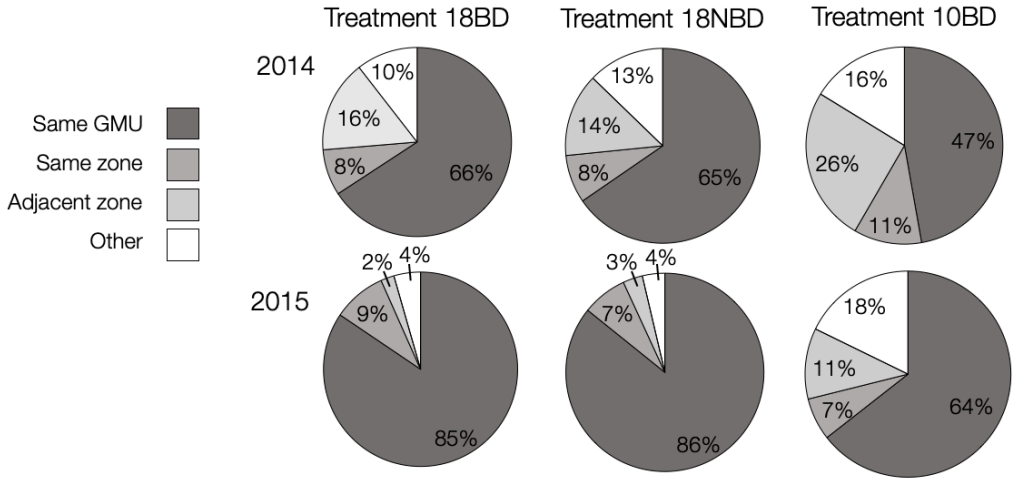

Figure 1.3: Percent of non-singleton sibling groups by categories of spatial distribution across cohorts and treatment types. Same GMU entails all members were reported within the same wolf management zone and the same game management unit. Same Zone depicts groups in which all members were within the same wolf management zone, but at least one member was not within the same game management unit. Groups categorized under Adjacent Zone contain at least one member that was not in the same wolf management zone as other putative siblings, but in which all members were reported to wolf management zones with shared borders. Groups categorized as Other contained at least one member that was in non-adjacent wolf management zone relative to other putative siblings.

Table 1.1: Parameter settings in the software Colony as altered by treatment, including the method used to calculate allele frequencies and the number of loci analyzed.

Treatment	Run ID	Run Length	Full-likelihood precision	Random number seed	Allele frequencies	Number of loci
18BD	1	Short	Low	Default	Known	18
	2	Short	High	Default	Known	18
	3	Short	High	Altered	Known	18
	4	Medium	High	Default	Known	18
	5	High	High	Default	Known	18
18NBD	1	Short	Low	Default	Unknown	18
	2	Short	High	Default	Unknown	18
	3	Short	High	Altered	Unknown	18
	4	Medium	High	Default	Unknown	18
	5	High	High	Default	Unknown	18
10BD	1	Short	Low	Default	Known	10
	2	Short	High	Default	Known	10
	3	Short	High	Altered	Known	10
	4	Medium	High	Default	Known	10
	5	High	High	Default	Known	10

Table 1.2: Mean frequency across runs of putative gray wolf litters by group size among harvested young of the year for both cohorts and all treatment types.

Year	Treatment	Number of group members				
		One	Two	Three	Four	Five
2014	18BD	14	20	8	4	0
	18NBD	6.2	21.4	9	2	2
	10BD	20.2	17.4	8.8	4.4	0
2015	18BD	16.2	18.8	8	3.8	0.2
	18NBD	18	20	8	3	1
	10BD	10	13	14	5	0

Table 1.3: Mean probabilities of inclusion and exclusion across all sibling groups of reconstructed configurations with a given treatment and cohort. *Indicates values significantly different from others within the same cohort based on Tukey HSD analysis and p value of < 0.05.

Year	Treatment	Inclusion	Exclusion
2014	18BD	0.95	0.56
	18NBD	0.97	0.57
	10BD	0.92*	0.40*
2015	18BD	0.97	0.48
	18NBD	0.97	0.49
	10BD	0.96	0.38*

Table 1.4: Rearrangement of sibling pairs by treatment and cohort. Mean Litter Count describes the mean total number of reconstructed sibling groups assigned within a given treatment and cohort, while Total Observed Pairs specifies unique individual sibling pairings within sibling groups. Rearranged Pairs describes the number of individual sibling pairings observed in at least one configuration that were not consistent across Colony runs within the same treatment.

Year	Treatment	Mean Litter Count	Total Observed Pairs	Rearranged Pairs
2014	18BD	52	126	0
	18NBD	53	131	9
	10BD	46.6	142	7
2015	18BD	63	205	0
	18NBD	63.8	210	13
	10BD	55	222	0

Table 1.5: Estimates of population parameters for 2014 and 2015 biological years, including (from left to right): minimum count of reproduction packs based on sibship reconstruction and estimates from other Idaho Department of Fish and Game (IDFG) monitoring efforts, total number of estimated packs within Idaho, IDFG's estimate of census size, total number of wolves harvested each year, the number of harvested young of the year genotyped, and the total number of young of the year used in sibling reconstructions, including individuals sampled from long-term study packs.

Cohort year	Minimum count (genetic sibship)	Minimum count (field observ.)	Est. packs in ID	Census estimate (IDFG)	Total harvested wolves	Harvested YOY genotyped	Total YOY* analyzed
2014	52	55	104	770	256	94	121
2015	63	69	108	786	256	101	162

Chapter 2: Comparing SNPs and Microsatellites for Sibling Reconstruction among Two Groups of Wolves

Heather Clendenin,[1,2] Jennifer Adams,[1] David Ausband,[3] Jim Hayden,[3] Paul Hohenlohe, [2] and Lisette Waits[1]

Departments of [1]Fish and Wildlife Sciences and [2]Biological Sciences: Bioinformatics and Computational Biology, University of Idaho; [3]Idaho Department of Fish and Game

INTRODUCTION

Advances in molecular biology and computational technologies have made it possible to assign genealogical relationships to organisms without reliable knowledge of breeding histories. As these tools have progressed, applications of relatedness analyses have expanded beyond forensics, epidemiology, and agriculture, and have come to provide insights into evolutionary and ecological processes and guide efforts to to promote genetic diversity in conservation programs (Artiles, Cobo, Benítez, Pérez, & Espinosa, 2015; Fernández et al., 2013; Hauser, Baird, Hilborn, Seeb, & Seeb, 2011; Sharp, Simeoni, & Hatchwell, 2008; Tokarska et al., 2009; Weinman, Solomon, & Rubenstein, 2014; Yu et al., 2015). As methods become both more affordable and more accessible, kinship studies based upon genetic markers are being used to verify observed pedigrees, maintain genetic diversity in captive-bred and managed wild populations, and permit monitoring of population metrics for managed wildlife and populations of conservation interest (Fienieg & Galbusera, 2015; Linløkken, Haugen, Mathew, Johansen, & Lien, 2016).

There are two categories of kinship analyses: relatedness estimation and relationship assignment. Relatedness estimations compare genetic data from individuals, and, based on assumptions of identity by descent, assign a continuous measure of relatedness based on shared fractions of individuals' genomes (Blouin, 2003; DeWoody, 2005). Relationship

assignments function similarly, but aim to identify putative genealogical relationship categories rather than relatedness indices. These methods can be employed in the construction of pedigrees, the identification of parentage, and the assignment of sibling groups. Which methods are applied to a given study can be affected by the ploidy and breeding systems of the organisms being studied, whether a population is predominantly inbred or outbred, as well as the consanguinity of individuals. Relationship assignments partition individuals into groups using likelihood estimations, set cover formulation, categorical or fractional designations, parentage inference, and sibship reconstruction (Anthony Almudevar & Field, 1999; Mary V Ashley et al., 2008; Painter, 1997; Thompson & Meagher, 1987; Jinliang Wang, 2004).

Sibship reconstruction attempts to assign individuals of the same generation to sibling groups (Anthony Almudevar & Anderson, 2012). Many of these methods use statistical likelihood, either groupwise or pairwise, and require knowledge of underlying allele frequencies (M. V. Ashley et al., 2009; Painter, 1997; Wang, 2004). A combinatorial approach based upon Mendelian properties, or a combination of both combinatorial and likelihood methods, may also be employed (Anthony Almudevar & Field, 1999; Mary V Ashley et al., 2008). These methods are not always sufficient when larger groups are analyzed; in these cases, combinatorial optimization, Markov chain Monte Carlo (MCMC) sampling, and model selection are preferable methodologies (Anthony Almudevar & Anderson, 2012).

Sibship assignments with or without parental information or known pedigree can be used as a tool for selective breeding in livestock populations and to maximize genetic diversity in captive bred populations (Fienieg & Galbusera, 2015). In the wild, assignment

of sibgroups can provide managers information on the effective number of breeders within a population (Linløkken et al., 2016). However, analysis results are subject to Type I errors of falsely assigning individuals to sibgroups, and Type II errors of falsely excluding individuals from true sibgroups. These assignment methods operate within given amounts of uncertainty tied to the number and genetic variability of loci used (DeWoody, 2005).

The predominant markers used in relatedness analyses have begun to shift in model organisms as well as in applications tied to industry and medicine. The transition from nuclear DNA microsatellite markers to single-nucleotide polymorphisms (SNPs) for relatedness estimates and kinship studies has been underway with species of economic interest, such as livestock and fisheries stocks (Fernández et al., 2013; Hauser et al., 2011; Yu et al., 2015). Some studies of evolutionary and ecological import have compared the power of SNPs and microsatellites within wild populations (Fabbri et al., 2012; Tokarska et al., 2009; Weinman et al., 2014). While previous predictions suggested that SNPs would be unlikely to become the predominant marker in non-model systems, these predictions may become anachronistic as hurdles are removed to high-throughput sequencing for SNP discovery and more applications in wild populations are observed (Andrew et al., 2013; Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Mary V Ashley et al., 2008; Benestan et al., 2016; Garner et al., 2016).

Monitoring demographic parameters is an important part of both appropriately managing wildlife species as well as a frequent component of regulatory requirements Previously, we explored the use of genetic sibling reconstruction of harvested young of the year as a method to provide an estimate of the minimum number of gray wolf litters within the state of Idaho (Clendenin et al., submitted). This effort was an attempt to minimize

resources required to estimate a minimum count of unique reproductive packs, but the results can also be combined with other demographic metrics to help establish a baseline rate of reproduction, depict the distribution of harvest pressure across packs, and serve as an index of the survival of young. Our initial forays into these methods built upon existing genetic monitoring based upon a suite of 18 microsatellite loci used to genotype individuals within this population since 2007. This approach generated robust results, with estimates comparable to those reported by local wildlife managers based upon field surveys, hunter reports, and other monitoring methods.

With the transition from genetic to genomics methods in mind, we were interested in evaluating the trade-offs between ongoing monitoring of reproductive packs through sibship reconstruction using existing microsatellite markers and the potential adoption of next-generation sequencing methods and SNPs. Though microsatellites are widely established in many wildlife monitoring programs, highly informative, and cheap to genotype, they are prone to genotyping error and the extensive training and subjectivity involved in their genotyping result in genotype calls that vary between, and even within, labs (DeWoody, 2005; Fernández et al., 2013; Pemberton, Slate, Bancroft, & Barrett, 1995). In contrast to microsatellites, SNPs are individually less informative, their initial discovery requires bionformatic expertise, and certain sequencing methods, such as RADseq, cost more per sample than microsatellite genotyping (Carroll et al., 2018; DeWoody, 2005). However, SNPs can be assayed by the hundreds or thousands, their genotyping can be standardized and automated, and the cost per marker is lower than microsatellites—and decreasing (Andrews, Barba, Russello, & Waits, 2018; Andrews et al., 2016). Also, while the burden of re-genotyping prior samples can be significant when transitioning a legacy dataset from

microsatellites to SNPs, SNPs provide an advantage over microsatellites when establishing new study systems given that no pre-existing genetic information is needed to identify loci.

To assess these trade-offs between marker choice for sibship analyses in populations with pre-existing microsatellite-based monitoring programs, we sought to answer three questions. First, how many SNP loci are required to provide sibling reconstruction results comparable to those based on the microsatellite loci currently used in monitoring? Secondly, how do measurements of accuracy and reliability compare between SNP-based sibship reconstructions and those based on microsatellite loci? How many SNPs are needed to substantially improve accuracy relative to the current suite of microsatellites? Finally, how do diversity characteristics and overall relatedness of a given population affect these outcomes?

We address these questions using both a cohort of wild, harvested gray wolf young of the year (YOY) from Idaho, as well as a separate group of wild, pedigreed, relatively inbred red wolf pups originating from a captive-bred population. In contrasting these two populations, we explore the relationship between levels of relatedness and amounts of informative genetic data required to correctly differentiating between siblings and related non-siblings.

METHODS

Red wolf sampling & microsatellite genotyping

We attempted to genotype a total of 101 blood and tissue samples from red wolves (53 red wolf pups born in 1999 and 48 from 2013) using microsatellites. These individuals arose from a captive-bred population that originated from 14 founding individuals, and a detailed

pedigree including these individuals has been reconstructed that provided *a priori* knowledge of true relationships (Bohling, Adams, & Waits, 2013). The 1999 samples were amplified in duplicate using 18 microsatellites as described in (Miller, Adams, & Waits, 2003). Among the 2013 cohort, multi-locus microsatellite genotypes at 17 loci were generated in two polymerase chain reactions. While the total number of loci was lower than the 1999 cohort, the individual genotyping rate exceeded 94% of 18 loci. Since this falls above the 90% genotyping success rate threshold set for inclusion, they are treated as if genotyped at 18 loci in downstream analyses.

For the 2013 cohort, the first multiplex contained 0.06 μM of CXX.377, 0.07 μM of CXX.172, CXX.173 and CXX.250, 0.13 μM of CXX.109, 0.16 μM of CXX.200, 0.20 μM of AHT121, 0.60 μM of AHT103, 0.71 μM of CXX.20, 1X Qiagen Multiplex PCR Kit Master Mix, 0.5X Q solution and 1 μl of DNA extract in a 7μl reaction (Holmes et al., 1995; Mellersh et al., 1997; Ostrander, Sprague, & Rine, 1993). The second multiplex contained 0.06 μM of FH2010, 0.07 μM of FH2062 and FH2054, 0.10 μM of FH2001, 0.16 μM of FH2145, 0.24 μM of FH2004, 0.36 μM of CXX.225, 0.80 μM of CXX403, 1X Qiagen Multiplex PCR Kit Master Mix, 0.5X Q solution and 1 μl of DNA extract in a 7 μl reaction (Mellersh et al., 1997; Ostrander, Mapa, Yee, & Rine, 1995; Ostrander et al., 1993). Reaction conditions for blood and tissue samples consisted of 13 cycles of 94 °C for 30 sec, 63 °C touchdown to 55 °C for 90 sec and 72 °C for 60 sec followed by 19 cycles of 94 °C for 30 sec, 55 °C for 90 sec and 72 °C for 60 sec followed by a final elongation of 60 °C for 10 minutes. Reaction conditions for the hair and bone samples were identical except the total number of cycles was 45 and 55 respectively. Blood and tissue samples were amplified a minimum of two times to account for possible allelic dropout. Hair and bone samples

were amplified up to four and six times respectively to account for possible allelic dropout. PCR products were run on a 3130xl Genetic Analyzer and visualized using Genemapper 5.0 (Applied Biosystems).

Gray wolf sampling & microsatellite genotyping

We obtained samples from 98 gray wolves determined to be YOY through tooth analysis. Samples were collected from May 2014 – April 2015, and include tissue from harvested wolves and other wolves of known mortality in management units throughout Idaho, as well as swabs from individuals captured for collaring. DNA extraction and genotyping at 18 microsatellite loci was conducted using methods documented in Clendenin et al., (submitted).

SNP discovery

SNP discovery for both red wolves and gray wolves was conducted using an updated RADseq protocol (Ali, Jeffres, & Miller, 2013). For red wolves, 86 of the original 101 samples contained sufficient quantities of genomic DNA, and these were prepared on a single 96-well plate with 10 sample duplicates. For gray wolves, the samples were equally divided between two plates with 52 samples each, including 6 sample duplicates. Standardized samples were digested with the restriction enzyme, Sbf-I HF. Libraries were sent to U.C. Berkeley genomics core facility for sequencing using a single high-output lane on an Illumina HiSeq 4000 machine with paired-end 150bp reads.

Reference-aligned SNP discovery using the dog genome, CanFam 3.1, was conducted using Stacks (Broeckx et al., 2014; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). We used a modified Mastretta-Yanes (2015) protocol to optimize Stacks

parameter settings by minimizing genotyping error between sample duplicates (Mastretta-Yanes et al., 2015). For red wolves, the parameters that differed from the default were the significance level for calling variant sites (0.001) and the significance level for calling genotypes (0.01). The optimized parameter for the gray wolves that differed from the default was setting a 0.01 chi square significance level for calling heterozygotes and homozygotes (Table 2.1).

Characterization of population diversity, relatedness, and effective population size

Estimates of genetic diversity, relatedness, and effective population size were generated for both red wolves and gray wolves using different marker data sets (Tables 2.2 and 2.3). GenAlex was used to report the number of alleles per locus, effective number of alleles per locus, Shannon's information index, unbiased expected heterozygosity, and observed heterozygosity (Peakall & Smouse, 2012). The R package, related, was used to obtain relatedness estimates using Milligan's dyadic likelihood method, as well as Wang's triadic likelihood method to address potential upward biases in the dyadic relatedness estimates (Milligan, 2003; Pew, Muir, Wang, & Frasier, 2015; Jinliang Wang, 2007, 2011). NeEstimator was used to provide a linkage disequilibrium-based estimate of effective population size as described in Waples and Do (2008), with each year also looked at individually among red wolves, and Colony2 was used to provide sibship-frequency based estimates of Ne (Do et al., 2014; Jones & Wang, 2010; Waples & Do, 2010).

Red wolf SNP filtering and Colony reconstruction

All SNPs were filtered using Plink to establish the maximum threshold of missing data per individual and per SNP, minor allele frequency threshold (MAF), Hardy-Weinberg

equilibrium, and linkage between SNPs. Different parameter combinations were tested systematically until subsequent Colony results were both consistent across ten replicates and matched the known pedigree (Table 2.1). Missing data filters reduced the number of samples to 75 individuals.

Among red wolves, sibling reconstructions were conducted with Colony using both microsatellite and the SNP genotypes. The microsatellite-based reconstructions used 100 individuals representing 25 sibling groups, with and without background population genetic data at the same 18 microsatellite loci based on the parental generations, and with and without assumed inbreeding. The SNP-based reconstructions were based on the 75 individuals that passed missing data thresholds described above. Plink filtering parameters were optimized and produced Colony results matched the known relationships from the pedigree across ten replicate Colony runs with at a set of 1509 SNPs (Table 2.1).

Additional subsets of these 1509 SNPs were generated by randomly thinning by 90%, 75%, 50%, 25%, and 10%, creating ten subsets per thinning category. Colony was run for each of these 50 new SNP subsets, and results were scored both by the accuracy of pairwise relationships (i.e. number of correctly identified true sibling pairs, number of true sibling pairs not identified, and number of non-siblings incorrectly assigned as sibling pairs) as well as by the total number of sibling groups generated (Figures 2.1-2.4).

Gray wolf SNP filtering and Colony reconstruction

While we were not able to evaluate genetic reconstructions relative to a known pedigree, we treated sibships from a previous study as a baseline of comparison for otherwise unknown relationships and litter counts among the harvested YOY (Clendenin et al., submitted).

Though the earlier study used the same suite of microsatellites, its results were consistent, known relationships within the dataset were accurately identified, and the total number of sibgroups was comparable to estimates based on other monitoring methods.

Two approaches were taken with respect to SNPs filtering for the gray wolves (Table 2.1). In one, we systematically varied Plink parameters and assessed Colony configurations based on their consistency across ten replicate runs and their similarity to the microsatellite-based results from our previous study. The other filtering approach applied variations on the same filters as those that produced the best results among the red wolf group. The SNPs isolated using the same parameters that were successful with the red wolves were thinned randomly by the same percentages as those used previously, while the other filtering group produced SNP sets with fewer loci which were therefore not thinned (Figure 2.5). Each final group of SNPs were evaluated based on ten replicate runs of Colony. Missing data filters reduced the number of samples to 50 individuals.

Sibling reconstructions based on microsatellites were reduced to the 50 individuals retained in the SNP data set. These were conducted both with and without population-level allele frequency estimates, and ten replicates were run for each variation of Colony parameter settings.

Geographical clustering of gray wolf siblings

As an additional assessment of accuracy of sibling assignments, we examined the geographical clustering of putative siblings within the best-performing SNP-based Colony configuration as well as the original microsatellite loci. All sibling groups of two or more for both the SNP-based and the original microsatellite-based configurations were categorized in

groups of increasing geographical distance, beginning with those including members all reported within the same game management unit (GMU), then respectively expanding to those with all members reported within the same wolf management zone (WMZ), within adjacent zones, or within non-adjacent zones that we could treat as spatial outliers (Figure 2.6). Though the designations weren't determined based on biological criteria, large geographical disparities among sibgroup members may be an indication of questionable or incorrect group assignments given that wolves at the age of sampling are unlikely to undertake long-distance dispersal. The categories of groups that differed between the SNP and microsatellite marker sets were compared, as well as the overall percentages of groups within each category.

RESULTS

Genotyping

Not all samples were successfully genotyped, and genotyping success rates varied between microsatellites and SNPs. Among the red wolves, all but one sample of the original 101 met the 90% genotyping success threshold at 18 microsatellite loci. Fewer samples were retained when genotyping with SNPs; 86 of 101 contained sufficient genomic DNA for RADseq library prep, and 71 samples passed through bioinformatics filtering. The 98 samples from the gray wolf cohort all contained sufficient DNA quantity for both genotyping at 18 microsatellite loci as well as library prep, but only 50 individuals passed through bioinformatics filtering.

Characterization of population diversity, relatedness, and effective population size

As expected, the number of alleles per locus and effective number of alleles per locus were higher for each microsatellite data set relative to the corresponding SNP datasets, but these values were also greater for each gray wolf marker type relative to the respective red wolf marker dataset (2.851 for red wolf microsatellites and 1.475 for SNPs; 4.726 for gray wolf microsatellites and 1.594 for SNPs). The Shannon information indices and estimates of heterozygosity followed the same pattern (Table 2.2).

Across all groups, the mean dyadic relatedness estimates were somewhat higher than the triadic estimates, but similar (Table 2.3). Estimates of mean relatedness based on microsatellite genotype data were higher relative to the corresponding SNP genotypes for both groups. Both estimates of mean relatedness among red wolves were higher than gray wolves for each genotype approach.

Estimates of effective population size followed a similar pattern, though the Ne estimates for red wolves were similar across marker types while the Ne estimates for gray wolves based on microsatellite data were much higher than those based on SNPs. All effective population size estimates for gray wolves were higher than those for red wolves, which was more pronounced when red wolf estimates were based on single cohorts (Figure 2.7).

Red wolf Colony reconstructions

The red wolf individuals genotyped using microsatellites represented 25 sibling groups from the known pedigree, and the subset of individuals successfully genotyped with SNP loci comprised 21 of those sibling groups. The Colony reconstructions using 18 microsatellite loci and the allele frequencies of the sampled cohort produced only 13 of the expected 25

sibling groups. Including allele frequency data from parental generations in the microsatellite-based sibling reconstructions increased the number of sibling groups to 20 of the expected 25.

After bioinformatic filtering with Plink, reconstructions based upon SNP genotypes at 1509 loci correctly identified all 21 of the expected subset of sibling groups as well as all individual relationships, and these sibling group assignments remained consistent across ten replicates. The Colony reconstructions using randomly thinned subsets of the selected 1509 SNP loci produced sibling groups with no errors with a fair degree of frequency at around 50% of the original filtered SNPs, using an average of 755 SNPs. While incorrect individual assignments were observed among the thinned SNP sets, errors in incorrectly assigned "false" siblings and incorrectly "missed" true siblings were nearly equal, meaning that total group counts were similar across all subsets, with as low as 10% of the original providing estimates of 21 total number of sibling groups (Figures 2.1-2.4).

Gray wolf Colony reconstructions

The subset of 50 gray wolf samples genotyped using SNPs represented 34 groups from our original full microsatellite-based Colony runs. Rerunning Colony using the same 18 microsatellite loci and the subset of 50 individuals consistently resulted in assignments with 33 sibgroups and no individual rearrangements. Two individuals previously assigned as singletons were paired in these new reconstructions.

In our first filtering approach with systematically varied parameters, the number of sibgroups with assignments consistent across ten replicates ranged from 19 groups using 101 loci up to 32 and 33 groups using 1029-3587 loci (Table 2.1). With our other filtering

approach that used the same filtering parameters optimized for the red wolf data set, 18,547 SNPs were produced among the gray wolf samples. Colony reconstructions using these SNPs produced 34 consistent sibgroups across 10 replicates, and this number of groups was retained when randomly thinning these SNP loci to as few as 50% of the original 18,547, or an average of 9251 loci (Figure 2.5).

Most of the 34 sibling groups generated by the SNP dataset corresponded to the 34 groups produced by our original microsatellite reconstructions. Between the original study and the new SNP-based reconstructions, 2 sibling groups were split and individuals from 2 other groups were joined. These differences persisted across all replicate runs.

Geographical clustering of putative gray wolf siblings

The subset of sibgroups from the original microsatellite reconstructions included 12 groups of 2 or more individuals. 9 of these groups contained siblings all within the same GMU, while there was one group in each of the other categories, i.e. within the same WMZ, within adjacent WMZ, or spatial outliers. The SNP-based reconstructions included 11 groups of 2 or more individuals. 8 groups had putative siblings all within the same GMU, 1 had siblings all within the same WMZ, 2 had siblings in adjacent WMZ, and no groups contained spatial outliers.

Most sibling groups had the same spatial clustering patterns across marker types However, the sibling groups that differed between the two marker data sets were clustered more closely among the SNP-based reconstructions (Figure 2.8). The microsatellite-based sibling group with individuals within adjacent zones was split into two singletons in the SNP-based reconstructions, and the group containing spatial outliers within the

microsatellite-based reconstructions was split into two groups in the SNP-based reconstructions—one containing individuals all within the same GMU and the other containing individuals within adjacent WMZ. The SNP-based reconstructions also joined two individuals in adjacent WMZ that had been designated as singletons in the microsatellite-based reconstructions. Additionally, no groups from the SNP-based reconstructions were categorized among the spatial outliers.

DISCUSSION

Sibling reconstruction has been used to study the evolution, ecology, and conservation of various wild populations by providing insights into dispersal, social dynamics, mating systems, and effective population size estimates (Ackerman et al., 2017; Dugdale, Macdonald, Pope, & Burke, 2007; Hansen & Jensen, 2005; Pope & Jha, 2017; Sharp et al., 2008). Our study presents valuable information about marker selection and the accuracy of results during an important time of transition of genetic methods. We set out to answer questions related to marker choice and sibling reconstruction, having previously demonstrated the reliability of genetic sibling reconstruction and its utility in meeting monitoring goals. Based on the results of our current study, it is evident that the best choice in a given study system will depend not just upon the markers selected and the question asked, but upon the characteristics of the population of interest.

While other studies have investigated differences in outcomes for analyses based on microsatellites or SNPs, many compare numbers of independent alleles or the effective number of alleles to provide optimal resolution without necessarily addressing the underlying mechanics of the given question within the populations of interest (Cappa, Klápště, Garcia, Villalba, & Marcucci Poltri, 2016; Hauser et al., 2011; Kaiser et al., 2017;

Linløkken et al., 2016; Weinman et al., 2014). Several kinship studies comparing markers have reported that the number of loci and their heterozygosity influences the power of a suite of markers more than the number of independent alleles (Hauser et al., 2011; Labuschagne, Nupen, Kotzé, Grobler, & Dalton, 2015; Morin, Luikart, Wayne, & the SNP workshop group, 2004). Though Wang (2019) has shown that even low quality samples with high genotyping error rates can produce reliable pedigrees, genotyping errors and missing data reduce the power of selected markers (Kaiser et al., 2017; Labuschagne et al., 2015; Jinliang Wang, 2019). The impact of population-level relatedness upon marker information required for sibling reconstruction and other kinship analyses has been covered obliquely in studies of livestock and other species with low genetic diversity, but has not been contrasted between two populations of closely related species with distinct differences in population parameters but similarities in vital rates and reproductive behaviors(Andrews, Adams, et al., 2018; Fernández et al., 2013; Fisher, Malthus, Walker, Corbett, & Spelman, 2009; Strucken et al., 2016; Tokarska et al., 2009; Trọng, van Bers, Crooijmans, Dibbits, & Komen, 2013).

Inbreeding poses a unique challenge to relatedness and kinship studies, as estimates of relatedness based on IBD inherently assume a historical reference generation in which all individuals are unrelated; inbreeding increases rates of allele sharing between relationship categories relative to those expected based on typical kinship coefficients (Blouin, 2003). Estimates can be further biased when allele frequencies are referenced from the sampled individuals under consideration rather than the broader population (J. Wang, 2014). Additionally, the unknown relatedness of founding members of captive-bred populations can further skew shared genomic fractions from those based upon typical assumptions (Herbinger, O'reilly, & Verspoor, 2006). Sets of genetic markers with greater information

content are required to provide resolution in relationship reconstruction among groups where these complications are observed, relative to groups that more closely resemble simplifying assumptions (Fernández et al., 2013; Sellars et al., 2014; Strucken et al., 2016; Tokarska et al., 2009; Trọng et al., 2013).

In evaluating prospective genetic monitoring protocols, the utility of marker sets with differing degrees of information content will vary by population. In our study, hundreds of SNP loci allowed for genetic sibling reconstructions that matched the known pedigree among the more inbred red wolves, while 18 microsatellite loci currently used for monitoring—even with parental allele frequency data—could not. The information content of the microsatellite loci was not sufficient to differentiate between siblings and closely-related non-siblings, while even random subsets of 10% (152 SNPs) of the full 1509 loci provided the same number of sibling groups. Alternately, among the gray wolves sampled, 18 microsatellite loci—with or without background data to estimate allele frequencies—consistently provided numbers of litters that closely corresponded to minimum estimates of reproductive packs obtained through field observations, hunter surveys, and other monitoring approaches (Clendenin et al., submitted). This set of microsatellites performed similarly to 1029-3587 SNP loci for the gray wolf population, and no further distinctions between groups were observed until the number of SNP loci was increased to over 9000 SNPs. With less overall relatedness among individuals in the population, sibling relationships could be satisfactorily identified using the current group of microsatellite loci used for long-term monitoring. While genetic measures of diversity and relatedness can have biases and shortcomings in representing population diversity, these estimates based on data from existing genetic monitoring programs can still be useful when considering a potential

transition from microsatellites to SNPs (Carroll et al., 2018; Väli, Einarsson, Waits, & Ellegren, 2008; J. Wang, 2014). Similarly, before implementing a new monitoring program, knowledge of recent demographic bottlenecks in wild populations or of a small number of founders within a captive-breeding program may help in marker selection.

Beyond the number of markers required, some interesting patterns were observed regarding the characteristics of the selected SNP loci. While SNP loci are biallelic, differences in MAF affect the information content of individual SNPs. SNPs with MAF approaching 0.5 are considered more informative than those with lower MAF, and often this is used as criteria for selection when establishing new SNP panels (Dussault & Boulding, 2018). Selecting SNPs with high MAF may be an effective strategy to decrease the number of loci required to obtain satisfactory relationship assignments and reduce genotyping costs per individual in some populations. However, inclusion of rare variants may improve the accuracy of relationship assignments, especially in highly inbred populations, by allowing for finer distinctions (Eynard, Windig, Leroy, van Binsbergen, & Calus, 2015). While missing data may also have played a contributing role, both perspectives on MAF and marker selection seem to have weight in our study; sibling reconstructions among the red wolves matched the known pedigree with a low MAF threshold (0.1), and while the gray wolf sibling reconstructions were the most discerning with filters mimicking those optimized for red wolves, performance comparable to microsatellites was possible with many fewer SNPs when a more moderate MAF threshold (0.3) was applied.

Looking to the future, there are considerations to weigh beyond the role of information content in the selection of markers as decisions are made about genetic monitoring protocols. Minimally-invasive samples with lower quality and quantities of

genomic DNA are often genotyped with microsatellites; while some next-generation sequencing methods, such as SNP arrays and target enrichment, can be suitable for such samples, other methods, such as the RADseq approach we employed, do not have the same success (Carroll et al., 2018). Costs of genotyping can vary across platforms and the accuracy of analyses is not strictly determined by marker type (Kraus et al., 2015; Puckett, 2017). In established, microsatellite-based programs with large population sizes and large genetic databases of samples, managers may find that the merits of retaining existing protocols outweigh the benefits of transitioning a legacy dataset to SNPs when the informativeness of the existing markers meet monitoring objectives and the costs to re-genotype years of existing samples is high. However, in populations with lower genetic diversity and higher levels of inbreeding or in populations that do not currently have a genetic monitoring program, SNPs may be preferable. In these cases SNPs have the advantage in that no existing genetic information is needed to select loci and genotyping success and per sample costs can be optimized using SNP chips and other technologies (Andrews, Barba, et al., 2018). When appropriately selected, SNPs can also allow for additional analyses, such as tests of adaptive differentiation, not possible with microsatellites. Additionally, even in systems where currently-employed microsatellite loci are adequate to answer questions of interest, transitioning to SNPs may be considered as broader platform shifts may result in fewer resources being developed for microsatellites and less support being available for related technologies and as pressure to include more data in publications increases (Carroll et al., 2018; Puckett, 2017).

LITERATURE CITED

Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C. A., … Campbell, M. R. (2017). Effective number of breeders from sibship reconstruction: empirical evaluations using hatchery steelhead. *Evolutionary Applications*, *10*(2), 146–160. https://doi.org/10.1111/eva.12433

Ali, O. A., Jeffres, C., & Miller, M. R. (2013). *RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping*. 16.

Almudevar, Anthony, & Anderson, E. C. (2012). A new version of PRT software for sibling groups reconstruction with comments regarding several issues in the sibling reconstruction problem. *Molecular Ecology Resources*, *12*(1), 164–178. https://doi.org/10.1111/j.1755-0998.2011.03061.x

Almudevar, Anthony, & Field, C. (1999). Estimation of Single-Generation Sibling Relationships Based on DNA Markers on JSTOR. Retrieved August 13, 2018, from https://ida.lib.uidaho.edu:2248/stable/1400594

Andrew, R. L., Bernatchez, L., Bonin, A., Buerkle, C. A., Carstens, B. C., Emerson, B. C., … Rieseberg, L. H. (2013). A road map for molecular ecology. *Molecular Ecology*, *22*(10), 2605–2626. https://doi.org/10.1111/mec.12319

Andrews, K. R., Adams, J. R., Cassirer, E. F., Plowright, R. K., Gardner, C., Dwire, M., … Waits, L. P. (2018). A bioinformatic pipeline for identifying informative SNP panels for parentage assignment from RADseq data. *Molecular Ecology Resources*, *18*(6), 1263–1281. https://doi.org/10.1111/1755-0998.12910

Andrews, K. R., Barba, M. D., Russello, M. A., & Waits, L. P. (2018). *Advances in Using Non-invasive, Archival, and Environmental Samples for Population Genomic Studies*. 1–37. https://doi.org/10.1007/13836_2018_45

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Artiles, A., Cobo, R., Benítez, L., Pérez, L., & Espinosa, G. (2015). Assessment of genetic variation and productive markers through four progenies of the first introduced stock of cultured shrimp *Penaeus (Litopenaeus) vannamei* in Cuba. *International Journal of Aquaculture*. https://doi.org/10.5376/ija.2015.05.0023

Ashley, M. V., Caballero, I. C., Chaovalitwongse, W., Dasgupta, B., Govindan, P., Sheikh, S. I., & Berger-Wolf, T. Y. (2009). KINALYZER, a computer program for reconstructing sibling groups. *Molecular Ecology Resources*, *9*(4), 1127–1131. https://doi.org/10.1111/j.1755-0998.2009.02562.x

Ashley, Mary V, Berger-Wolf, T. Y., Caballero, I. C., Chaovalitwongse, W., DasGupta, B., & Sheikh, S. I. (2008). *Full Sibling Reconstruction in Wild Populations From Microsatellite Genetic Markers*. 29.

Benestan, L. M., Ferchaud, A.-L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., … Luikart, G. (2016). Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular Ecology*, *25*(13), 2967–2977. https://doi.org/10.1111/mec.13647

Blouin, M. S. (2003). DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution*, *18*(10), 503–511. https://doi.org/10.1016/S0169-5347(03)00225-8

Bohling, J. H., Adams, J. R., & Waits, L. P. (2013). Evaluating the ability of Bayesian clustering methods to detect hybridization and introgression using an empirical red wolf data set. *Molecular Ecology*, *22*(1), 74–86. https://doi.org/10.1111/mec.12109

Broeckx, B. J. G., Coopman, F., Verhoeven, G. E. C., Bavegems, V., De Keulenaer, S., De Meester, E., … Deforce, D. (2014). Development and performance of a targeted whole exome sequencing enrichment kit for the dog (Canis Familiaris Build 3.1). *Scientific Reports*, *4*, 5597. https://doi.org/10.1038/srep05597

Cappa, E. P., Klápště, J., Garcia, M. N., Villalba, P. V., & Marcucci Poltri, S. N. (2016). SSRs, SNPs and DArTs comparison on estimation of relatedness and genetic parameters' precision from a small half-sib sample population of Eucalyptus grandis. *Molecular Breeding*, *36*(7), 97. https://doi.org/10.1007/s11032-016-0522-7

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L., & Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary Applications*, *11*(7), 1094–1119. https://doi.org/10.1111/eva.12600

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

DeWoody, J. A. (2005). Molecular Approaches to the Study of Parentage, Relatedness, and Fitness: Practical Applications for Wild Animals. *The Journal of Wildlife Management*, *69*(4), 1400–1418.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Dugdale, H. L., Macdonald, D. W., Pope, L. C., & Burke, T. (2007). Polygynandry, extra-group paternity and multiple-paternity litters in European badger (Meles meles) social groups. *Molecular Ecology*, *16*(24), 5294–5306. https://doi.org/10.1111/j.1365-294X.2007.03571.x

Dussault, F. M., & Boulding, E. G. (2018). Effect of minor allele frequency on the number of single nucleotide polymorphisms needed for accurate parentage assignment: A methodology illustrated using Atlantic salmon. *Aquaculture Research*, *49*(3), 1368–1372. https://doi.org/10.1111/are.13566

Eynard, S. E., Windig, J. J., Leroy, G., van Binsbergen, R., & Calus, M. P. (2015). The effect of rare alleles on estimated genomic relationships from whole genome sequence data. *BMC Genetics*, *16*(1), 24. https://doi.org/10.1186/s12863-015-0185-0

Fabbri, E., Caniglia, R., Mucci, N., Thomsen, H. P., Krag, K., Pertoldi, C., … Randi, E. (2012). Comparison of single nucleotide polymorphisms and microsatellites in non-invasive genetic monitoring of a wolf population. *Archives of Biological Sciences*, *64*(1), 321–335. https://doi.org/10.2298/ABS1201321F

Fernández, M. E., Goszczynski, D. E., Lirón, J. P., Villegas-Castagnasso, E. E., Carino, M. H., Ripoli, M. V., … Giovambattista, G. (2013). Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genetics and Molecular Biology*, *36*(2), 185–191. https://doi.org/10.1590/S1415-47572013000200008

Fienieg, E., & Galbusera, P. (2015). *The use and integration of molecular DNA information in conservation breeding programmes: a review*. 6.

Fisher, P. J., Malthus, B., Walker, M. C., Corbett, G., & Spelman, R. J. (2009). The number of single nucleotide polymorphisms and on-farm data required for whole-herd parentage testing in dairy cattle herds. *Journal of Dairy Science*, *92*(1), 369–374. https://doi.org/10.3168/jds.2008-1086

Garner, B. A., Hand, B. K., Amish, S. J., Bernatchez, L., Foster, J. T., Miller, K. M., … Luikart, G. (2016). Genomics in Conservation: Case Studies and Bridging the Gap between Data and Application. *Trends in Ecology & Evolution*, *31*(2), 81–83. https://doi.org/10.1016/j.tree.2015.10.009

Hansen, M. M., & Jensen, L. F. (2005). Sibship within samples of brown trout *(Salmo trutta)* and implications for supportive breeding. *Conservation Genetics*, *6*(2), 297–305. https://doi.org/10.1007/s10592-004-7827-5

Hauser, L., Baird, M., Hilborn, R., Seeb, L. W., & Seeb, J. E. (2011). An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (Oncorhynchus nerka) population. *Molecular Ecology Resources*, *11 Suppl 1*, 150–161. https://doi.org/10.1111/j.1755-0998.2010.02961.x

Herbinger, C. M., O'reilly, P. T., & Verspoor, E. (2006). Unravelling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. *Molecular Ecology*, *15*(8), 2261–2275. https://doi.org/10.1111/j.1365-294X.2006.02923.x

Holmes, N. G., Dickens, H. F., Parker, H. L., Binns, M. M., Mellersh, C. S., & Sampson, J. (1995). Eighteen canine microsatellites. *Animal Genetics*, *26*, 131–132.

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

Kraus, R. H. S., vonHoldt, B., Cocchiararo, B., Harms, V., Bayerl, H., Kühn, R., … Nowak, C. (2015). A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf monitoring in Europe based on noninvasively collected samples. *Molecular Ecology Resources*, *15*(2), 295–305. https://doi.org/10.1111/1755-0998.12307

Labuschagne, C., Nupen, L., Kotzé, A., Grobler, P. J., & Dalton, D. L. (2015). Assessment of microsatellite and SNP markers for parentage assignment in ex situ African Penguin (Spheniscus demersus) populations. *Ecology and Evolution*, *5*(19), 4389–4399. https://doi.org/10.1002/ece3.1600

Linløkken, A. N., Haugen, T. O., Mathew, P. K., Johansen, W., & Lien, S. (2016). Comparing estimates of number of breeders Nb based on microsatellites and single nucleotide polymorphism of three groups of brown trout (Salmo trutta L.). *Fisheries Management and Ecology*, *23*(2), 152–160. https://doi.org/10.1111/fme.12169

Mellersh, C. S., Langston, A. A., Acland, G. M., Fleming, M. A., Ray, K., Wiegand, N. A., … Ostrander, E. A. (1997). A linkage map of the canine genome. *Genomics*, *46*, 326–336.

Miller, C. R., Adams, J. R., & Waits, L. P. (2003). Pedigree-based assignment tests for reversing coyote (Canis latrans) introgression into the wild red wolf (Canis rufus) population. *Molecular Ecology*, *12*(12), 3287–3301. https://doi.org/10.1046/j.1365-294X.2003.02003.x

Milligan, B. G. (2003). *Maximum-Likelihood Estimation of Relatedness*. 16.

Morin, P. A., Luikart, G., Wayne, R. K., & the SNP workshop group. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, *19*(4), 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Ostrander, E. A., Mapa, F. A., Yee, M., & Rine, J. (1995). One hundried and one new simple sequence repeat-based markers for the canine genome. *Mammalian Genome*, *6*, 192–195.

Ostrander, E. A., Sprague, G. F., & Rine, J. (1993). Identification and characterization of dinucleotide repeat (CA)n markers for genetic mapping in dog. *Genomics*, *16*, 207–213.

Painter, I. (1997). Sibship Reconstruction without Parental Information. *Journal of Agricultural, Biological, and Environmental Statistics*, *2*(2), 212–229. https://doi.org/10.2307/1400404

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pemberton, J. M., Slate, J., Bancroft, D. R., & Barrett, J. A. (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, *4*(2), 249–252. https://doi.org/10.1111/j.1365-294X.1995.tb00214.x

Pew, J., Muir, P. H., Wang, J., & Frasier, T. R. (2015). related: an R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*, *15*(3), 557–561. https://doi.org/10.1111/1755-0998.12323

Pope, N. S., & Jha, S. (2017). Inferring the foraging ranges of social bees from sibling genotypes sampled across discrete locations. *Conservation Genetics*, *18*(3), 645–658. https://doi.org/10.1007/s10592-017-0941-y

Puckett, E. E. (2017). Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions. *Conservation Genetics Resources*, *9*(2), 289–304. https://doi.org/10.1007/s12686-016-0643-7

Sellars, M. J., Dierens, L., McWilliam, S., Little, B., Murphy, B., Coman, G. J., … Henshall, J. (2014). Comparison of microsatellite and SNP DNA markers for pedigree assignment in Black Tiger shrimp, Penaeus monodon. *Aquaculture Research*, *45*(3), 417–426. https://doi.org/10.1111/j.1365-2109.2012.03243.x

Sharp, S. P., Simeoni, M., & Hatchwell, B. J. (2008). Dispersal of sibling coalitions promotes helping among immigrants in a cooperatively breeding bird. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*(1647), 2125–2130. https://doi.org/10.1098/rspb.2008.0398

Strucken, E. M., Lee, S. H., Lee, H. K., Song, K. D., Gibson, J. P., & Gondro, C. (2016). How many markers are enough? Factors influencing parentage testing in different livestock populations. *Journal of Animal Breeding and Genetics*, *133*(1), 13–23. https://doi.org/10.1111/jbg.12179

Thompson, E. A., & Meagher, T. R. (1987). Parental and Sib Likelihoods in Genealogy Reconstruction. *Biometrics*, *43*(3), 585–600. https://doi.org/10.2307/2531997

Tokarska, M., Marshall, T., Kowalczyk, R., Wójcik, J. M., Pertoldi, C., Kristensen, T. N., … Bendixen, C. (2009). Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. *Heredity*, *103*(4), 326–332. https://doi.org/10.1038/hdy.2009.73

Trọng, T. Q., van Bers, N., Crooijmans, R., Dibbits, B., & Komen, H. (2013). A comparison of microsatellites and SNPs in parental assignment in the GIFT strain of Nile tilapia (Oreochromis niloticus): The power of exclusion. *Aquaculture*, *388–391*, 14–23. https://doi.org/10.1016/j.aquaculture.2013.01.004

Väli, Ü., Einarsson, A., Waits, L., & Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*, *17*(17), 3808–3817. https://doi.org/10.1111/j.1365-294X.2008.03876.x

Wang, Jinliang. (2004). Sibship Reconstruction From Genetic Data With Typing Errors. *Genetics*, *166*(4), 1963–1979. https://doi.org/10.1534/genetics.166.4.1963

Wang, J. (2014). Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. *Journal of Evolutionary Biology*, *27*(3), 518–530. https://doi.org/10.1111/jeb.12315

Wang, Jinliang. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research*, *89*(03). https://doi.org/10.1017/S0016672307008798

Wang, Jinliang. (2011). coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, *11*(1), 141–145. https://doi.org/10.1111/j.1755-0998.2010.02885.x

Wang, Jinliang. (2019). Pedigree reconstruction from poor quality genotype data. *Heredity*, 1. https://doi.org/10.1038/s41437-018-0178-7

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, *3*(3), 244–262. https://doi.org/10.1111/j.1752-4571.2009.00104.x

Weinman, L. R., Solomon, J. W., & Rubenstein, D. R. (2014). A comparison of SNP and microsatellite markers for analysis of parentage and kinship in a cooperatively breeding bird. *Molecular Ecology Resources*, n/a-n/a. https://doi.org/10.1111/1755-0998.12330

Yu, G. C., Tang, Q. Z., Long, K. R., Che, T. D., Li, M. Z., & Shuai, S. R. (2015). Effectiveness of microsatellite and single nucleotide polymorphism markers for parentage analysis in European domestic pigs. *Genetics and Molecular Research*, *14*(1), 1362–1370. https://doi.org/10.4238/2015.February.13.15

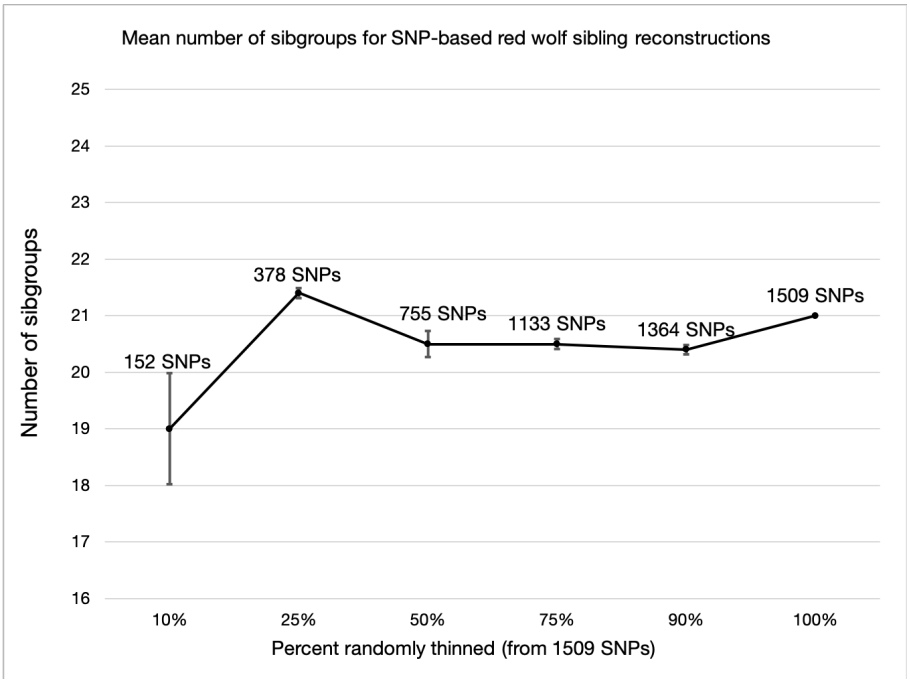

Figure 2.1: Count of red wolf litters/sibgroups using SNPs randomly thinned from best-performing 1509, averaged over 10 subsets per "percent randomly thinned" category.

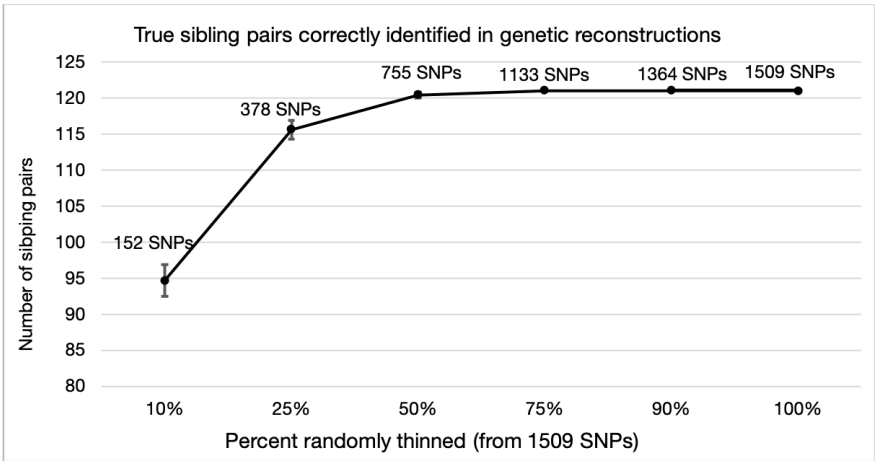

Figure 2.2: Rates of correct pairwise assignments of red wolves using SNPs randomly thinned from best-performing 1509, averaged over 10 subsets per "Percent thinned" category. The average number of correct sibling pairs are the true siblings correctly identified in Colony reconstructions.

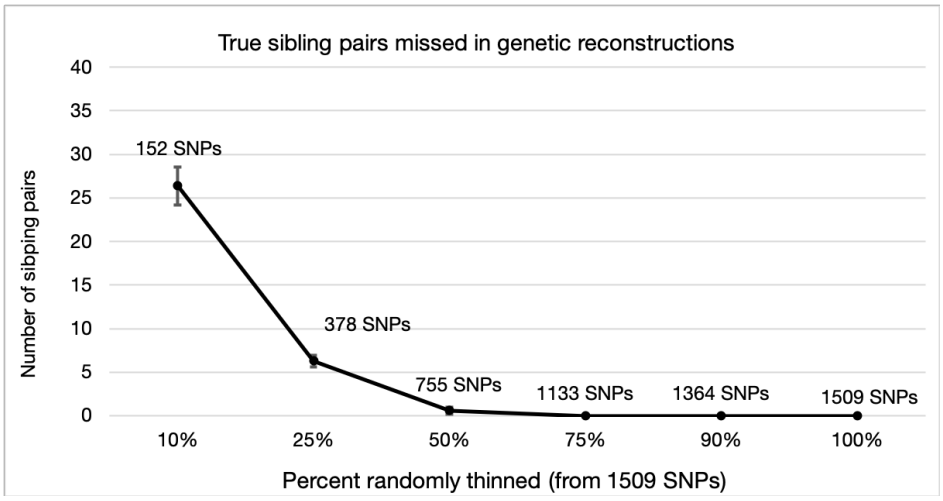

Figure 2.3: Rates of missed true red wolf sibling pairs using SNPs randomly thinned from best-performing 1509, averaged over 10 subsets per "Percent thinned" category. Missed true sibling pairs are true siblings not included in the same sibling group in Colony reconstructions.

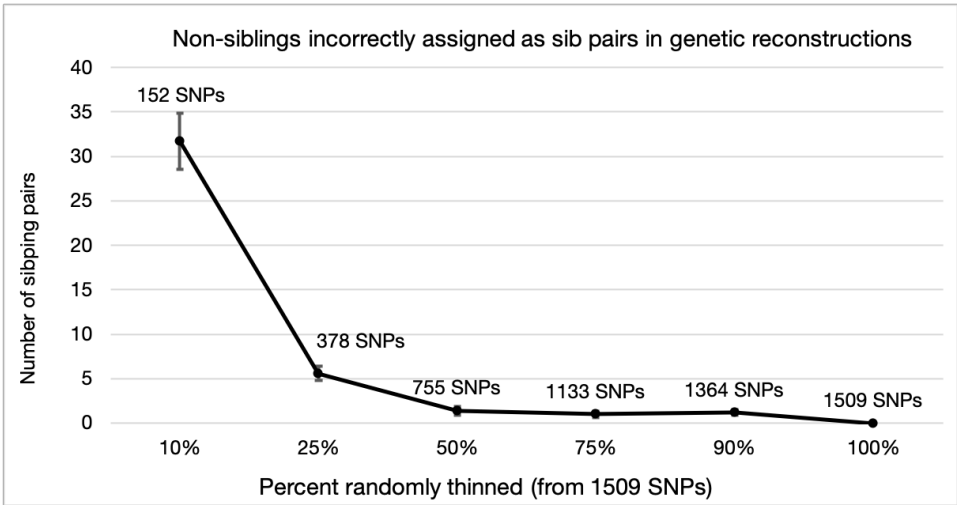

Figure 2.4: Rates of incorrectly assigned red wolf sibling pairs using SNPs randomly thinned from best-performing 1509, averaged over 10 subsets per "Percent thinned" category. Incorrectly assigned sibling pairs are known non-siblings included in the same sibling group in Colony reconstructions.

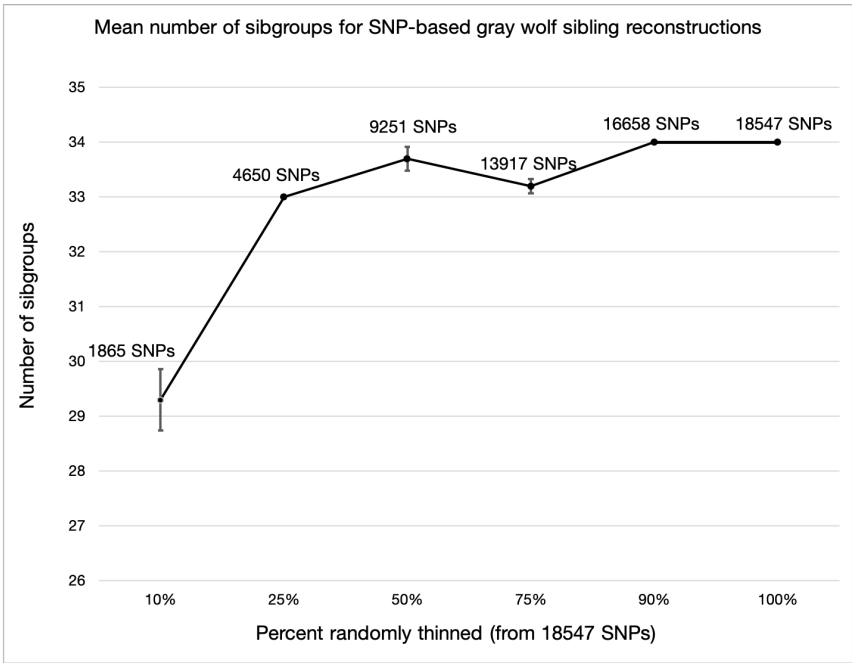

Figure 2.5: Count of gray wolf litters using SNPs randomly thinned from best-performing 18457, averaged over 10 subsets per "percent randomly thinned" category.

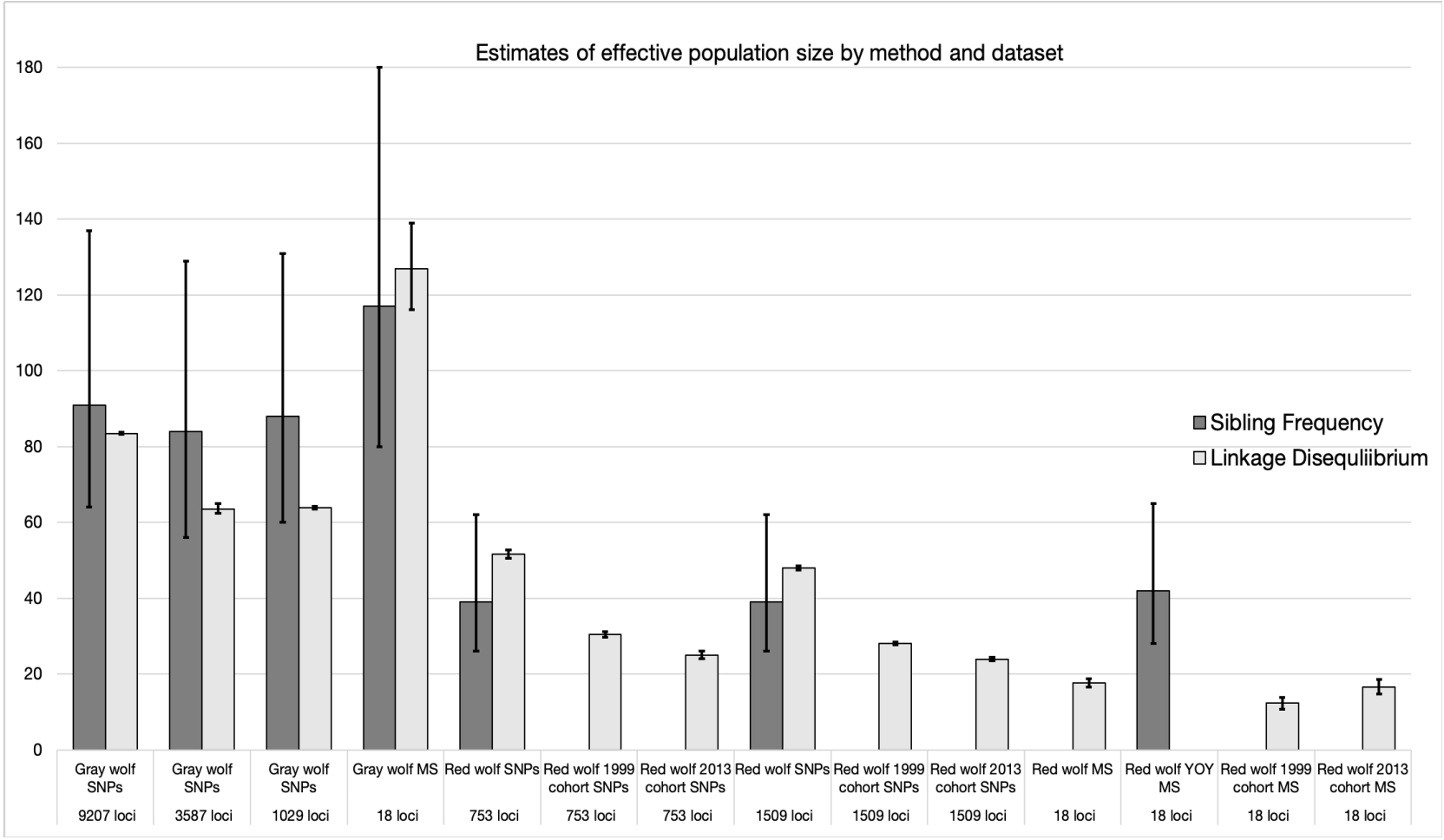

Figure 2.6: Estimates of effective population size by method and dataset. Bars correspond with estimates of Ne with 95% confidence intervals. Not all methods were implemented for each dataset.

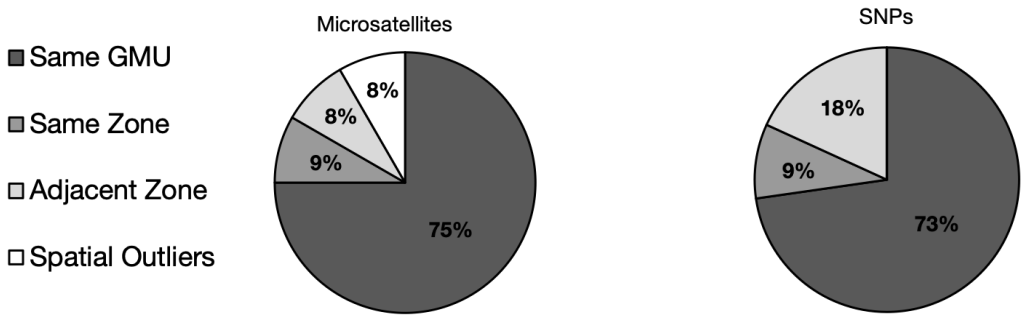

Figure 2.7: Geographical clustering of putative gray wolf sibling groups with two or more members, based on best-performing Colony reconstructions for each marker type. Categories are designated in order of increasing geographical distance. "Same GMU" describes sibling groups with all members reported within the same game management unit (GMU). "Same Zone" describes sibling groups with all members reported within the same wolf management zone (WMZ). "Adjacent Zone" describes sibling groups with all members reported within adjacent zones. "Spatial Outliers" describe sibling groups that contain members within non-adjacent zones.

Table 2.1: Count of litters and accuracy of Colony reconstructions by species and marker set. "RW" indicates the information describes the red wolf analyses, while "GW" indicates gray wolves. The "NBD" method refers to reconstructions with no background data used to estimate population-level allele frequencies, while "BD" indicates these data were used. "Pedigree matching" describes Plink filtering parameters selected by subsequent Colony reconstructions that matched the known red wolf pedigree. "Subset MS"describes reconstructions using microsatellites with individual samples subset to match those retained in SNP-based reconstructions. "RW opt. params" describes the dataset of gray wolf SNPs obtained through use of the Plink filtering parameters optimized for the red wolf dataset. "Syst Var" describes datasets of gray wolf SNPs obtained by systematically varying Plink parameters to obtain subsequent Colony reconstructions that were close to the original microsatellite-based reconstructions with the full set of individuals and also consistent over 10 replicate runs. "Min/ind" refers to the Pink parameter that establishes a threshold for missing genotype data per individual sampled, "MAF" refers to the minor allele frequency threshold, and "Min/SNP" refers to the threshold for missing genotype data per SNP. "#Litters obs./exp." details the total number of litters averaged over 10 replicate Colony runs over the expected total based on either pedigree (for red wolves) or the benchmark established from previous microsatellite-based reconstructions (gray wolves).

Species	Method	Marker type	#Loci	Min/ind	MAF	Min/SNP	#Litters obs./exp.
RW	NBD	MS	18	NA	NA	NA	13/25
	BD	MS	18	NA	NA	NA	20/25
	Pedigree matching	SNP	1509	0.96	0.1	45	21/21
GW	Subset MS	MS	18	NA	NA	NA	33/34
	RW opt. params	SNP	18547	0.96	0.1	45	34/34
	Syst Var	SNP	101	0.954	0.1	20	19/34
	Syst Var	SNP	2349	0.954	0.1	30	32/34
	Syst Var	SNP	1029	0.954	0.2	30	32/34
	Syst Var	SNP	6504	0.954	0.2	40	33/34
	Syst Var	SNP	3587	0.954	0.3	40	33/34

Table 2.2: Diversity metrics with standard errors (SE) from GenAlex results based on microsatellites (MS) and SNPs from red wolves (RW) and gray wolves (GW). Includes the number of alleles per locus (Na), effective number of alleles per locus(Ne), Shannon's information index (I), observed heterozygosity (Ho), and unbiased expected heterozygosity (uHe).

	Na	**Ne**	**I**	**Ho**	**uHe**
RW SNPs	2.000 (SE +-0.000)	1.475 (SE +- 0.006)	0.477 (SE +- 0.003)	0.250 (SE +- 0.003)	0.309 (SE +- 0.003)
RW MS	6.056 (SE +- 0.408)	2.851 (SE +- 0.181)	1.231 (SE +- 0.064)	0.652 (SE +- 0.032)	0.627 (SE +- 0.027)
GW SNPs	2.000 (SE +- 0.000)	1.594 (SE +- 0.003)	0.535 (SE +- 0.001)	0.272 (SE +- 0.001)	0.361 (SE +- 0.001)
GW MS	8.000 (SE +-0.676)	4.726 (SE +- 0.447)	1.642 (SE +- 0.095)	0.717 (SE +- 0.029)	0.756 (SE +- 0.026)

Table 2.3: Mean population relatedness based on microsatellites (MS) and SNPs from red wolves (RW) and gray wolves (GW). The triadic method uses uses Wang's 2007 triad-based maximum likelihood estimator, and the dyadic method uses Milligan's 2003 dyad-based maximum likelihood estimator.

Marker set	Triadic	Dyadic
RW SNPs	0.056	0.058
RW MS	0.105	0.120
GW SNPs	0.031	0.031
GW MS	0.070	0.081

Appendix

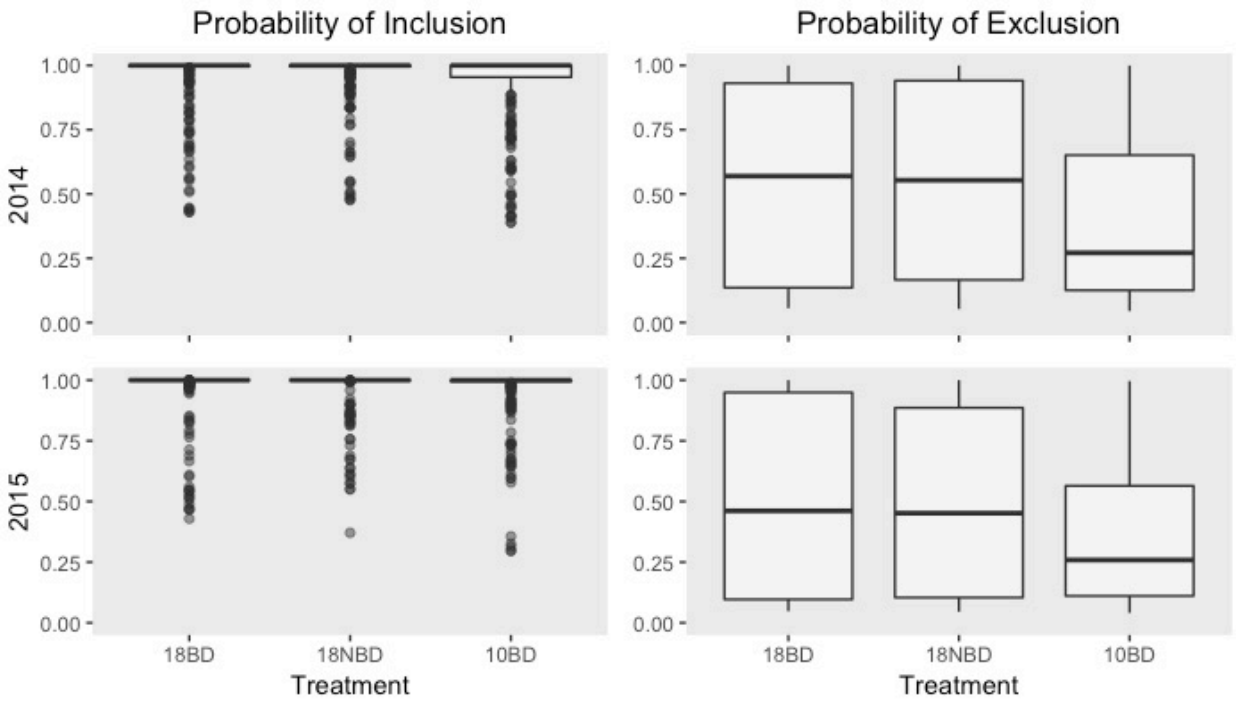

Figure S1.1: Distribution of mean probabilities of inclusion and exclusion for sibling groups across Colony treatments, 2014 and 2015.

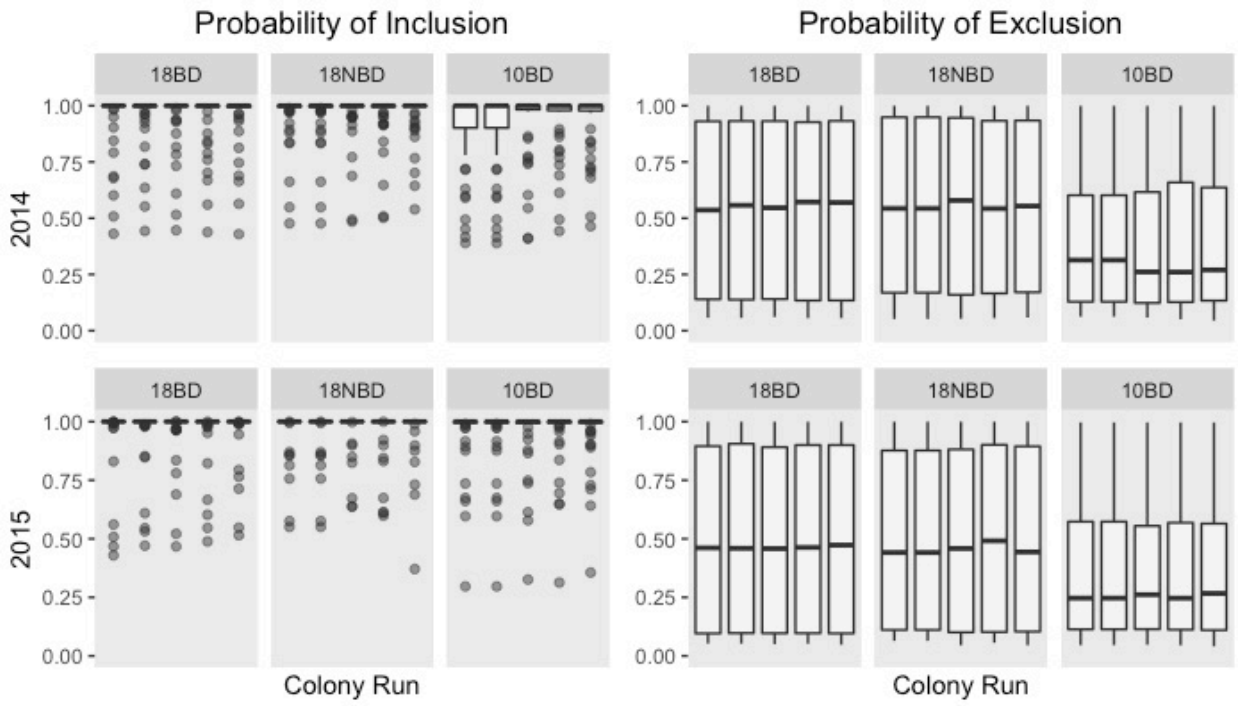

Figure S1.2: Distribution of mean probabilities of inclusion and exclusion for sibling groups across Colony runs and treatments, 2014 and 2015.

Table S1.1: Loci and allelic diversity for 18 microsatellites used in this study. Abbreviations include: A for number of alleles; H_e and H_o for expected and observed heterozygosity, respectively.

Marker	A	Ho	He
103	8	0.77	0.736
109	6	0.76	0.771
121	8	0.78	0.807
172	4	0.42	0.430
200	8	0.45	0.624
2004	12	0.83	0.856
2010	4	0.58	0.610
250	8	0.68	0.790
377	8	0.77	0.784
CXX119	10	0.81	0.774
CXX173	6	0.71	0.773
FH2001	8	0.76	0.761
FH2054	11	0.73	0.790
FH2088	9	0.67	0.671
FH2137	13	0.77	0.751
FH2611	13	0.86	0.885
FH2670	17	0.88	0.879
FH3725	13	0.76	0.816