# Applications of Effective Population Size in Conservation Biology Through an Experimental Study and Quantitative Review

Shannon H. Clarke

A Thesis in

The Department of Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science (Biology) at

Concordia University

Montreal, Quebec, Canada

January 2022

**CONCORDIA UNIVERSITY**

School of Graduate Studies

This is to certify that the thesis prepared

By: Shannon Clarke

Entitled: Applications of effective population size in conservation biology through an experimental study and quantitative review

And submitted in partial fulfillment of the requirements for the degree of

**Master of Science (Biology)**

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee

| | |
|---|---|
| _____ *Dr. Dylan Fraser* | Supervisor/Chair |
| _____ *Dr. I. Ferguson* | External Examiner |
| _____ *Dr. C. Ziter* | Examiner |
| _____ *Dr. S. Dayanandan* | Examiner |
| _____ *Dr. James Grant* | Co-Supervisor |

Approved by _____
*Dr. R. Weladji – Graduate Program Director*

_____ 2022 _____
*Dean of Faculty*

## GENERAL ABSTRACT

Applications of effective population size in conservation biology through an experimental study and quantitative review

Shannon H. Clarke

Effective population size ($N_e$) is a key parameter in conservation biology, as it influences the degree of genetic drift and adaptive potential of a population. Although $N_e$ is widely reported in individual studies, it is not well integrated into monitoring programs, and there is still much to learn about how it is impacted by human-induced stressors. My thesis attempted to bridge these gaps through (1) an experiment assessing how $N_e$ changes within populations exposed to size-selective harvest, and (2) a quantitative review summarizing $N_e$ estimates across and within taxa in wild populations. I showed experimentally that in populations of brook trout (*Salvelinus fontinalis*) exposed to size-selective harvest, change in $N_e$ is buffered through density-dependent genetic compensation. Therefore, in the short-term, $N_e$ can be resilient to harvest even with substantial decreases in population size. My review also showed that $N_e$ varies between taxonomic groups in wild populations, and that a high human footprint is associated with lower $N_e$, especially in amphibians and mammals. There were two broad conclusions from this work. Firstly, many wild populations fall below key conservation thresholds ($N_e$ of 50 or 500), and I discuss the caveats of using conservation thresholds based on broad generalizations across taxa that vary in life history traits and adaptations. Secondly, I emphasize the importance of integrating genetic and demographic factors in monitoring and risk assessments. Overall, these conclusions can help guide the integration of $N_e$ into research, monitoring, and policy.

**Acknowledgements**

First and foremost, I would like to thank my supervisors, Dr. Dylan Fraser and Dr. James Grant, without whom this work would not have been possible. Thank you for your guidance, support, and many hours spent reading my writing. I would also like to thank the members of my committee, Dr. Carly Ziter and Dr. Selvadurai Dayadandan, for their assistance and feedback during my MSc. This research was funded by an NSERC Strategic Project Grant to Dylan Fraser, and an NSERC Canada Graduate Scholarship and a GRIL Collaborative Partnership Award to Shannon Clarke.

This project would not have been possible without the support of Parks Canada for permits, logistical support, and use of their facilities. I would like to thank Brent Brookes and Dylan Glaser for their significant contributions to the field research associated with Chapter 1, along with Justin Budyk, Hannah Postma, Giulio Navarroli, Natalie Dupont, Julien Beaulieu, and other field crew members for their help in sampling. Thank you all for the great memories spent in the beautiful Rocky Mountains. This research was greatly facilitated by Gregory McCracken and Dr. Daniel Ruzzante from Dalhousie University, who sequenced the brook trout microsatellites. I would also like to thank Robert Carson and Ashlee Prevost for their advice and help with DNA extraction at Concordia University.

This work would also not have been possible without the tremendous support I received during the screening and data collection for Chapter 2. Thank you to Brian Gallagher, Sarah Salisbury, Sozos Michaelides, Ramela Koumrouyan, Elizabeth Lawrence, and Jean-Michel Matte. Special thanks to Elizabeth for completing the mapping in ArcGIS, and to Jean-Michel for his advice and troubleshooting with R and mixed modelling.

I would also like to thank all the members of the Fraser Lab for their feedback and support during my MSc, and to all the friends that made my time at Concordia as memorable as it was. To my parents, Paul and Linda, thank you for your unconditional love and support, and genuine interest in my research, and for always believing in me. Finally, thank you to my partner, Robby Earl, for supporting me unconditionally over the last three years, even if it meant I would be away for several months at a time during field work.

**Contribution of Authors**

*Chapter 1*

Study concept and design: Clarke, Humphries, Grant, Fraser

Acquisition of data: Clarke, McCracken,

Analysis and interpretation of data: Clarke

Drafting of manuscript: Clarke

Critical revision: Clarke, McCracken, Ruzzante, Grant, Fraser

*Chapter 2*

Study concept and design: Clarke, Fraser

Acquisition of data: Clarke, Lawrence, Matte, Gallagher, Salisbury, Michaelides, Koumrouyan

Analysis and interpretation of data: Clarke, Lawrence, Matte

Drafting of manuscript: Clarke

Critical revision: Clarke, Grant, Fraser

## Table of Contents

**List of Figures**

intervals. Groups with no shared letters are statistically different from one another. Models were generated using the non-replicated dataset.

**Figure 2.5**: Relationship between effective population size ($N_e$) and human footprint across taxonomic groups. Shaded areas represent 95% confidence intervals.

**Figure 2.6**: Effective to census size ratio ($N_e/N_c$ and $N_b/N_c$) across taxonomic groups, accounting for differences between $N_e$ and $N_b$. Shaded green bars represent the 95% confidence intervals. Groups with no shared letters are statistically different from one another.

**List of Tables**

**List of Supplementary Material**

**Figure S1**: Flow diagram of the review process, modified by O'dea et al. (2021)

**Figure S2**: Flow diagram of the filtering process, modified by O'dea et al. (2021). *Some records had both spatial and temporal replication

**Figure S3**: Regression between $N_e$ estimates produced from the Linkage Disequilibrium (LD) method, and the Bayesian Computation method. Sample size for the analysis was N = 416.

**Figure S4**: Regression between $N_e$ estimates produced from the Linkage Disequilibrium (LD) method, and the Sibship Frequency method. Sample size for the analysis was N = 386.

**Figure S5**: Number of studies with $N_e$ or $N_b$ estimates published per year. Studies from 2020 were not included as the literature search was conducted partway through the year and is not representative of all articles published

**Figure S6**: Number of estimates of $N_e$ or $N_b$ sampled per year in the final dataset. There were 957 estimates with no sampling year reported.

**Figure S7**: Mean effective number of breeders ($N_b$) across taxonomic groups, accounting for differences in marker type. Shaded green bars represent 95% confidence intervals. *The upper CI for plants and marine fish are not contained within the plot. Groups with no shared letters are statistically different from one another

**Figure S8**: Global maps showing the taxa-specific median $N_e$ value in each 500×500km grid cell. Data is projected with the world Behrmann projection

## General Introduction

Human activity is having an increasing impact on wild populations worldwide, from climate change, to habitat fragmentation, and over-exploitation (Bellard et al., 2012; Rosser & Mainka, 2002; Wilson et al., 2016). Evaluating the risk to populations and developing effective management strategies has therefore become a primary focus of conservation biology (Burgman et al., 1993; Burgman & Yemshanov, 2013; Frankham, 2005; Harwood, 2000). Additionally, there has been an increased focus on incorporating genetic analyses into risk assessment and management regimes, as inbreeding and loss of genetic diversity can play a key role in extinction of wild populations (Frankham, 2003; Ottewell et al., 2016; Ralls et al., 2018). The effective population size ($N_e$) is an important parameter in conservation genetics, as it gives a measure of both inbreeding and the rate of genetic drift in populations (Waples, 2002). $N_e$ is defined as the size of an idealized population that experiences genetic drift at the same rate as the actual population (Crow & Kimura, 1970). In an ideal population, $N_e$ will be equal to the actual number of individuals (the census size, $N_c$), however in wild populations this is rarely the case (Palstra & Fraser, 2012; Ruzzante et al., 2016). $N_e$ is often much lower than $N_c$ in wild populations, due to factors such as fluctuations in population size, variance in reproductive success between individuals, or an unequal sex ratio (Franklin, 1980). Therefore, $N_e$ can be a useful tool in evaluating a population's risk of extinction.

Recent developments in molecular technologies and software have allowed for much greater ease of estimating $N_e$, providing even more rationale for its use in conservation biology. There are two general approaches to estimating $N_e$ in populations: demographic, or ecological, methods, and genetic methods. Demographic methods involve using information about a population's life history, such as the sex ratio or variance in family size, to estimate $N_e$ (Nunney

& Elam, 1994). Genetic methods use information from genetic markers, like microsatellites or Single Nucleotide Polymorphisms (SNPs), to track inbreeding or genetic drift and generate an $N_e$ estimate (Wang, 2005). Recently, there has been an increase in the use and development of single-sample genetic methods such as the Linkage Disequilibrium method or the Sibship Frequency method, due to their ease of use (Wang, 2016). Whereas previously researchers needed in-depth information about the life-history of a population (demographic method), or needed multiple temporal estimates separated by a sufficient time lag (temporal genetic method), single-sample estimators allow for estimation from data within a single sampling season. Indeed, this ease of use allowed for a drastic increase in the number of single-sample estimates published from 1990 to 2011, and this trend has likely continued over the past decade (Palstra & Fraser, 2012). Due to their increased accessibility and widespread use, single-sample genetic estimators of $N_e$ are a key tool that can be used in conservation biology.

In order for $N_e$ to be a useful tool in wild population management, it is important to establish what values of $N_e$ are cause for concern in wild populations. Because $N_e$ reflects the level of inbreeding in a population, and can be an indicator of extinction risk, $N_e$ is often lower in populations of conservation concern compared to non-listed populations (Palstra & Ruzzante, 2008). Thresholds of minimum $N_e$, known as the 50/500 rule, have been created as guidelines to ensure wild populations are viable in both the short- and long-term (Franklin, 1980). According to the 50/500 rule, a population must have a minimum $N_e$ of 50 to reduce the likelihood for extinction in the short-term due to inbreeding depression, and a minimum $N_e$ of 500 to retain enough genetic variation to allow for adaptation in the long-term. More recent works have even suggested increasing the minimum thresholds to 100/1000, though there is some debate over whether this is necessary (Frankham et al., 2014; García-Dorado, 2015; Jamieson & Allendorf,

2012). Not only can $N_e$ be used as an indicator of extinction risk, but it can also be used to measure the impacts of human activities on populations. Previous studies have used $N_e$ or $N_e/N_c$ to monitor populations exposed to pollution (e.g. Durrant et al. 2011), habitat fragmentation (e.g. Alò and Turner 2005; Chavez-Pesqueira et al. 2014), and overfishing (e.g. Therkildsen et al. 2010; Mirimin et al. 2016), though many of these are observational rather than experimental. Experimental studies in wild populations using $N_e$ as an indicator of population health are less common but can provide key insights into the impacts of human activities on $N_e$. Effective population size can therefore be used to address empirical questions regarding human influences, and to assess the risk of populations to extinction.

In the first chapter of this thesis, I assessed how size-selective harvest influenced effective population size and its ratio with census size, through a multi-lake experimental study in alpine populations of brook trout (*Salvelinus fontinalis*). Using a Before-After Control-Impact (BACI) design over three years, I compared $N_e$ in lakes exposed to intensive harvest (~0.59 harvest rate) to non-harvested control lakes. I predicted that the census size would decrease much more rapidly than the effective size, due to competitive release from removing the largest individuals (i.e. genetic compensation). The practical application of this work, alongside the targeted experimental design, showed how effective size can be used as a tool for monitoring change in exploited populations, and highlighted the importance of using experimental (rather than observational) studies to explore relationships.

In Chapter 2, I conducted a quantitative review of single-sample estimators of $N_e$ in wild populations, assessing taxonomic differences and evaluating the impacts of human activity. Using systematic review approaches, I compiled a database of 4972 effective size estimates from 712 peer-reviewed articles, and geo-referenced all populations to generate estimates of the

Global Human Footprint (Venter et al., 2016). Specifically, I assessed whether taxonomic groups differed in their effective size, whether different taxa were more or less likely to fall below the 50/500 thresholds, and how $N_e$ was related to human footprint. Based on differences in life history traits, I expected there to be taxonomic differences in $N_e$ and the likelihood of reaching 50/500. I also hypothesized that $N_e$ would be negatively correlated with human footprint, but that some taxa might be more impacted than others, e.g., those most impacted by human activities like fragmentation or pollution. The broad scope of this chapter demonstrated the use of $N_e$ in a comprehensive analysis that relates directly to risk assessments and can inform changes to listing criteria.

My thesis highlights how $N_e$ can be used in a variety of applications and contributes new knowledge about both small- and large-scale trends in $N_e$. Chapter 1 revealed how effective size responds in the short-term after large reductions in census size. Chapter 2 demonstrated how previous reviews reporting $N_e$ or $N_e/N_c$ across taxa were not representative for certain taxonomic groups. Overall, these results have important implications by showing how $N_e$ can be used in fisheries monitoring, and by providing guidance on which taxa should be prioritized for risk assessment based on the disproportionate impact of human activities.

**Chapter 1: Experimental harvesting of wild brook trout populations reveals genetic compensation**

Shannon H. Clarke, Gregory R. McCracken, Shelley Humphries, Daniel E. Ruzzante, James W.A. Grant, and Dylan J. Fraser

*This chapter has been submitted for review to the journal Evolutionary Applications, and is formatted to meet the journal guidelines

**Abstract**

Sustainable fisheries management benefits from integrating demographic and genetic considerations into fishery assessments, as both play a role in determining harvest yields and population persistence. This is especially important in populations subject to size-selective harvest, which has the potential to result in significant demographic, life-history, and genetic changes. One important parameter used to assess genetic changes and predict future population viability is the effective population size ($N_e$), which is often compared with the census size ($N_c$) to integrate genetic and demographic changes. Estimating $N_e$ in wild populations is difficult, so an alternate measure from a single reproductive cohort (also known as $N_b$, or the effective number of breeders), can instead be used to compare with $N_c$. We investigated harvest-induced changes in $\widehat{N}_b$ (referring to an estimate, rather than the parameter $N_b$) for introduced brook trout populations (*Salvelinus fontinalis*) in alpine lakes from western Canada. Three populations were subject to size-selective harvesting, while three control populations experienced no harvest. Changes in $\widehat{N}_b$ and $\widehat{N}_c$ were compared between control and harvested populations over three years of harvest. The $\widehat{N}_c$ decreased consistently across all harvested populations (an average decline of 60.8%) but fluctuated in control populations. While no consistent changes in $\widehat{N}_b$ between control or harvest populations were detected, one harvest population experienced a significant decrease in $\widehat{N}_b$ of 63.2%. We also found evidence of genetic compensation in the harvested populations, where density-dependent processes buffered changes in $N_b$ such that $\widehat{N}_b/\widehat{N}_c$ increased at smaller population sizes (a ten-fold increase on average). While the results of this study suggest that effective size in harvested populations may be resilient to considerable changes in $N_c$ in the short-term, it is still important to monitor exploited populations to assess the risk of inbreeding and ensure their long-term survival.

## Introduction

An ongoing concern in fisheries and wildlife management is ensuring the sustainability of both commercial and recreational harvest. Selective pressures introduced through harvest can drive evolutionary and ecological changes, and potentially reduce harvest yields and wild population persistence (Fraser 2013; Kuparinen and Festa-Bianchet 2017). One of the most common harvest strategies in fisheries is size-selective harvest, where the largest individuals are targeted, generating selection on body size (Stokes and Law 2000; Law 2000). Size-selective harvest often results in considerable life history changes, including earlier maturation, reduction in body size, and reduced fecundity (Kuparinen and Merilä 2007a; Fenberg and Roy 2008). Demographic processes also play a role in determining the effects of size-selective harvesting on exploited populations; processes such as density-dependence, competitive release, or Allee effects may influence how populations respond to or recover from harvest (Zipkin et al. 2008; Kuparinen et al. 2014; Gobin et al. 2016). Recent studies have also highlighted the importance of tracking genetic change in exploited populations in order to predict long-term viability and ensure sustainable management (Allendorf et al. 2008; Ovenden et al. 2015). Therefore, monitoring both genetic and demographic processes can play an important role in managing populations exploited by size-selective harvesting.

One important genetic parameter for evaluating threats to wild populations is the effective population size ($N_e$); it integrates both life history and genetic changes and can track inbreeding and genetic drift in populations (Waples 2002). In managed populations, $N_e$ is often considered alongside the demographic population size, or census size ($N_c$), in order to combine both demographic changes and genetic processes into a single metric. The ratio of $N_e/N_c$ in an idealized population should be ~1, because all individuals should be contributing equally to

successive generations; however, in natural populations, $\widehat{N}_e$ is often much lower than $\widehat{N}_c$ (Palstra and Ruzzante 2008; Palstra and Fraser 2012; Ruzzante et al. 2016). The factors that drive $\widehat{N}_e$ below $\widehat{N}_c$ include unequal sex ratios, variance in reproductive success, and fluctuations in $N_c$ (Franklin 1980); all of which may occur in exploited populations. Estimating $N_e$ in wild populations, however, can be difficult due to the need for demographic data from a species' entire lifespan, or the need for multiple temporal genetic samples (Waples et al. 2014). Instead, many studies use the effective number of breeders ($N_b$), which is a measure of effective size for a single reproductive season, making it a much more accessible parameter than $N_e$ (Ferchaud et al. 2016). $N_b$ is closely linked with $N_e$ through life-history traits (Waples et al. 2013) and is also often compared alongside $N_c$ to evaluate both demographic and genetic changes in managed populations (Palstra and Fraser 2012).

The ratio between $\widehat{N}_b$ and $\widehat{N}_c$ may differ between populations of the same species due to differences in population size, environmental variation, or different life history characteristics. Larger populations generally have a higher $\widehat{N}_b$ than smaller populations (Yates et al. 2017); however, density-dependent effects on reproduction often lead to very large populations having a smaller $\widehat{N}_b/\widehat{N}_c$ ratio (Bernos and Fraser 2016; Ferchaud et al. 2016). In addition, environmental variation may influence $\widehat{N}_b/\widehat{N}_c$ in species with specific breeding or habitat requirements (Shrimpton and Heath 2003; Whiteley et al. 2015). Finally, life history characteristics such as body size, age at maturity, and mating systems can also impact $\widehat{N}_b/\widehat{N}_c$ (Jones and Hutchings 2001; Lee et al. 2011; Waples and Antao 2014). Most research on these topics has been observational – not experimental – and done spatially across populations, not temporally within populations. Although the general trends found in observational studies may be representative of how demographic and genetic processes act within populations, experimental studies are still

necessary to fully understand how populations may respond temporally to disturbances such as harvest.

Over the long term, intensive harvest is expected to decrease both $N_b$ (by reducing the number of potential adults available to breed, e.g., Diaz et al. 2000) and $N_c$ (by increasing mortality rates). However, $\widehat{N}_c$ is likely to decrease more rapidly than $\widehat{N}_b$, leading to an increase in $\widehat{N}_b/\widehat{N}_c$. According to Kuparinen et al.'s (2016) model, $\widehat{N}_c$ decreased proportionally more than $\widehat{N}_e$ during simulated harvest of cod stocks, and in some cases $\widehat{N}_e$ actually increased (the impacts on $N_e$ and $N_b$ are expected to be similar). Indeed, many demographic factors moderate changes in $N_b$ and may even lead to increases, especially in the short-term. For example, density-dependent effects may result in genetic compensation (i.e., where $N_b/N_c$ increases) at small population sizes due to a decrease in competition for reproductive opportunities resulting in a decrease in the variance in reproductive success (Ardren and Kapuscinski 2003; Bernos and Fraser 2016). Size-selectivity during harvest can also moderate changes to $N_b$ through competitive release, since in some populations spawning is dominated by the largest size-classes (Blanchfield et al. 2003; Anderson et al. 2010). Furthermore, harvest may alter $N_b$ if it leads to changes in the sex ratio within a population (a skewed sex ratio will tend to result in lower $N_b$). Size-selective harvest can alter sex ratios if the sexes differ in size, or if one sex (often male) is more vulnerable to harvest than the other (Allendorf and Hard 2009; Kendall and Quinn 2013). Therefore, $N_b$ and $N_b/N_c$ are expected to be impacted by size-selective harvesting through a variety of demographic and genetic processes.

Iteroparous salmonid fishes are ideal study species to observe changes in $N_b/N_c$ from size-selective harvest. $\widehat{N}_b/\widehat{N}_c$ ratios in salmonid species are well documented in the literature and can vary from almost zero to upwards of 0.5 (Ferchaud et al. 2016; Bernos et al. 2018). Salmonids

are also characterized by specific spawning habitat, and intense competition for reproductive opportunities, therefore they are very likely to be affected by density-dependent processes (Ardren and Kapuscinski 2003). In brook trout (*Salvelinus fontinalis*), females compete for sites of varying quality on breeding grounds (Blanchfield and Ridgway 2005), and breeding is often skewed towards larger males (Blanchfield et al. 2003). Finally, iteroparous salmonids such as brook trout have short generation times (e.g. Letcher et al. 2007; Gray et al. 2014), which allows for shorter-term studies on size-selective harvesting that can span changes in $\widehat{N}_b/\widehat{N}_c$ across an entire generation.

This study addresses gaps in the literature by examining how genetic ($N_b$) and demographic ($N_c$) variables respond to size-selective harvest. Specifically, we asked how size-selective harvesting influenced $\widehat{N}_b$ and $\widehat{N}_b/\widehat{N}_c$ in populations of brook trout in the Rocky Mountains. We hypothesized that $\widehat{N}_c$ would decrease after successive years of intensive harvest while $\widehat{N}_b$ would either be unchanged or decrease at a slower rate than $\widehat{N}_c$ due to competitive release from removing the largest individuals in the population, with $\widehat{N}_b/\widehat{N}_c$ therefore increasing at lower population sizes (i.e. genetic compensation). We also expected that secondary effects of harvest, such as changes to the sex ratio or reductions in body size, would have an impact on $\widehat{N}_b$, and that there may be population-specific responses to harvest based on the specific demographics or life history of each population.

## Materials and Methods

### *Study System*

The brook trout populations of interest inhabited six alpine lakes across two national parks in western Canada: Banff and Kootenay National Parks (Figure 1.1). These lakes ranged in

elevation from 1470m to 2400m, and were chosen due to their physical isolation, small size (ranging from 1.7 to 18 ha.), fish communities dominated by brook trout, and limited inlet or outlet expanse (Table S1). Previous work in this system showed no difference in neutral genetic diversity between populations (due to their common stocking source), but showed some adaptive genetic differentiation (Brookes et al., 2022). Although most lakes were isolated, two of them (Margaret and McNair) could potentially be open to seasonal gene flow from adjacent populations through outlet waterfalls during extreme weather events (Thompson and Rahel 1998; Adams et al. 2000). The populations in these lakes were assumed to be in an equilibrium state due to little or no exploitation before this project (a few lakes were subject to minimal catch-and-release recreational fishing). We subjected three of these lakes to size-selective harvesting over three successive years, with an average harvest rate of 0.59 (i.e., 59% of all adult fish were harvested, beginning with the largest fish and slowly decreasing the size of harvested fish; Table S2). This harvest rate is close to those seen in some fisheries in marine and freshwater environments (Allen et al. 2008; Thomas et al. 2009; Topping and Szedlmayer 2013). The other three lakes were left as non-harvested controls but were monitored over the same time span.

*Census Size Estimation and Experimental Harvest of Lakes*

The $\widehat{N}_c$ of brook trout in each lake was obtained through mark-recapture methods during the summers of 2017-2019. We captured fish of all size classes using fyke nets and inserted BioMark HPT8 pre-loaded Passive Integrated Transponder (PIT) tags into adults >80mm (those less than 80mm were categorized as juveniles), therefore our $\widehat{N}_c$ estimates corresponded with the adult population. We also measured fork length (in mm) for each fish captured. On each successive visit, the number of tagged fish and untagged fish was recorded, all the PIT tag IDs of

tagged fish were noted, and all new fish were also inserted with a PIT tag. Fyke nets were checked daily to minimize stress for fish, and marking took, on average, 13 days (ranging from 7 to 20), depending on the habitat, catchability, and efficiency of capture methods.

We used the Schnabel method to estimate $\widehat{N}_c$ using multiple recapture events. This method is widely used in fisheries science and has several assumptions including: the population is closed (no emigration or immigration), the fish retain their tags (PIT tag retention in salmonids is generally >95%; (Gries and Letcher 2002; Dare 2003; Ostrand et al. 2011), and there is an equal probability of catching all fish (Huggins and Chao 2005). The Schnabel estimate was generated on the mark-recapture data using the FSA package (fisheries stock assessment, version 0.8.30) in R Statistical Software (Ogle et al. 2019; R Core Team 2019). This package generated an estimate of $\widehat{N}_c$ along with upper and lower 95% confidence intervals associated with the estimate.

Once $\widehat{N}_c$ was estimated, ~9% of the fish from each lake (ranging from 4% to 15%) were harvested using multi-mesh gill nets to target all size classes. These fish were lethally sampled in order to obtain standardized size, age, and sex data, which allowed us to contrast the life histories and demographics spatially between populations and temporally within populations. In the harvest lakes, the size structure (determined from length measurements from the multi-mesh gill nets) was used to determine which mesh size of gill net was needed to target the largest individuals, while allowing the smallest ones to pass through. These gill nets were then placed in the harvest lakes and checked every 1-3 days. The gill nets were only removed when the targeted harvest rate was reached. During this time the control lakes were left untouched. This harvest period occurred in late summer and early fall, before individuals were able to breed, therefore the

$\widehat{N}_b$ calculated in the following year should relate to only those individuals left after the harvest was completed (see Figure 1.2).

*Tissue Collection and Microsatellite Sequencing*

To calculate $\widehat{N}_b$ annually in each lake, young-of-the-year (YOY) trout were sampled in 2017, 2019 and 2020 using a backpack electrofisher. The shoreline of each lake was electrofished, targeting fish under ~80mm. These fish were measured (in mm), a small clip was taken from their caudal fin and placed in an Eppendorf tube with 95% ethanol, and the fish were released. If there were insufficient numbers of YOY along the shoreline, then inlets and outlets were fished to supplement the sample size. YOY samples were taken from as many locations along the shoreline or inlet/outlets as possible, to ensure the $\widehat{N}_b$ calculated from these samples was representative of the whole population (rather than a single brood). Prior to sample analysis, we compared the lengths of individuals to a length-age curve to ensure we only sampled YOY, and no 1+ individuals were included in the analyses.

When evaluating $\widehat{N}_b/\widehat{N}_c$ ratios it is important that the $\widehat{N}_b$ and $\widehat{N}_c$ values are correctly linked – that is, that the $\widehat{N}_b$ estimate from a cohort is matched with the $\widehat{N}_c$ estimate from the parental generation that produced that cohort (Fig 1.2). The YOY samples collected in 2017 (i.e. the 0+ cohort) would therefore correspond to "pre-harvest", as they would have been produced by the parental generation in 2016. Because we did not have census data from 2016, the 2017 YOY samples were paired with 2017 census data, as this census data should also correspond to "pre-harvest" (in other non-exploited populations, Bernos and Fraser (2016) compared "incorrectly" linked $\widehat{N}_b/\widehat{N}_c$ from the same year and found similar patterns as in correctly linked ratios). Because the harvest period occurred prior to spawning, when calculating the $\widehat{N}_b/\widehat{N}_c$ ratio

for 2019 and 2020 we accounted for any individuals removed during harvest (for treatment lakes) or during age and maturity determination (in both control and treatment lakes).

DNA was extracted from fin clips of 1248 individuals using a modified Chelex protocol (see Yue and Orban 2005). The quality and quantity of the DNA extracts was assessed via visualization on 1% agarose gel stained with SYBR Safe (Thermo Fisher Scientific, Waltham, MA), and a Qubit fluorometer (Invitrogen). These extracts were sequenced at 33 microsatellite loci (from Ruzzante et al. 2019) using an Illumina MiSeq Sequencer along with the MEGASAT software and pipeline (Zhan et al. 2017). Further details on the molecular protocol and loci can be found in Appendix A.

The microsatellite scoring was verified manually using the depth vs size histograms produced by MEGASAT. Microchecker (v2.2.3; Van Oosterhout et al. 2004) was used to test for potential null alleles, large allele dropout, and accidental scoring of stutter bands. Hardy-Weinberg Exact (HWE) tests for heterozygote deficiency and excess, as well as tests for linkage disequilibrium (LD) between loci were performed in GenePop (v4.7.5; Rousset 2008). Observed and expected heterozygosity and mean number of alleles were calculated in Arlequin (v3.5.2.2; Excoffier and Lischer 2010).

*Estimation of Effective Number of Breeders*

Estimates of $\widehat{N}_b$ were generated using the Linkage Disequilibrium (LD) method implemented in NeEstimatorV2 (Do et al. 2014). The LD method is a single-sample estimator first developed by Hill (1981) that uses non-random associations of alleles at different loci (i.e. deviations from the expected genotype frequency based on random distribution) to estimate random drift and therefore $N_e$ or $N_b$. Here, because all individuals are from the same cohort, our

estimates correspond with $N_b$. The LD method implemented in NeEstimator v2 is a bias-corrected version, that reduces downward bias when sample size is smaller than the true effective size (Waples 2006). We used an allele frequency cutoff of 0.02, a value that balances precision and bias across sample sizes (Waples and Do 2008), and 95% confidence intervals were generated using a jackknife approach.

*Statistical Analyses*

All statistical analyses were performed in R statistical software version 4.0.5 (R Core Team 2021). To test whether $\widehat{N}_b$ was reduced through harvest, we used a linear mixed model with time (i.e., year) and treatment (harvest/control) as the predictor variables, $\widehat{N}_b$ as the response variable, and lake (i.e., each population) as a random factor (random intercept). Similarly, a linear mixed model was performed using $\widehat{N}_b/\widehat{N}_c$ as the response variable and time and treatment as the predictor variables and lake as a random factor in order to test whether genetic compensation occurred with harvest. For both of these models, if the harvest treatment had an effect on $\widehat{N}_b$ or $\widehat{N}_b/\widehat{N}_c$, we would expect to find a significant interaction between treatment and time, i.e., that over three years of successive harvest, the harvest lakes responded differently than the control lakes. All models were run using the *lme* function in the *nlme* package (version 3.1; Pinheiro et al. 2017) and were tested for normality through visualization of a qqplot (using the *qqnorm* function in the *stats* package; R Core Team 2021), as well as for homoscedasticity by plotting the normalized residuals against the fitted values. The response variable in both models were log-transformed to allow the data to approximate a normal distribution. Post-hoc analyses were done using the *lsmeans* package (Lenth 2018) to test for differences between treatments. We also ran a linear mixed model with census size ($\widehat{N}_c$) as the

response variable, using the same predictors and random variables as above. No transformations were needed for this model, but all other methods were the same as with the models for $\widehat{N}_b$ and $\widehat{N}_b/\widehat{N}_c$.

During size-selective harvest, because life history traits and changes in $\widehat{N}_c$ were expected to influence $\widehat{N}_b$, we also tested how $\widehat{N}_b$ changed with $\widehat{N}_c$, sex ratio, and variance in body size using mixed models. Due to the small sample size of our dataset (n = 18), we could not create a full model with all of the predictor variables, so instead we created three competing models based on a priori hypotheses and compared their AIC to see which best explained the data. The first model included $\widehat{N}_c$, treatment, and their interaction as predictors; the second model included the coefficient of variation (CV) of body length (i.e. the standard deviation divided by the average), treatment, and their interaction as predictors; and the third model included the sex ratio (males / females), treatment, and their interaction as predictors. All models also included lake as a random intercept variable (random intercept models always had a lower AIC than random intercept and slope). After testing the normality of the model residuals, none of the three models conformed to the assumption of normality, therefore generalized linear mixed models were run with a gamma distribution. A gamma distribution was chosen because $N_b$ values are continuous (i.e. non-integer), and are always positive (a negative estimate indicates that there was insufficient sample size to accurately estimate $N_b$). GLMMs were run using the *glmer* function in the *lme4* package (Version 1.1-26; Bates et al. 2020), using Adaptive Gauss-Hermite Quadrature (nAGQ = 0) in order to allow model convergence; using nAGQ = 0 does not fully account for the randomness of the random effects, but is often quite comparable with the default of nAGQ = 1, using the Laplace approximation, as it does little to lower the deviance (Bates 2011; Stegmann et al. 2018). After comparing the AIC from the three models, we also performed backward model

selection, dropping terms from each of the three models to see whether simpler models were better at explaining the data.

## Results

### *Census Size Estimates*

Census size estimates ranged from 195 individuals (McNair Lake 2018) to 2,800 individuals (Olive Lake 2017) (Table S3). All of the harvest lakes experienced a decrease in census size from 2017 to 2019 (average decline of 60.8%, ranging from 55.5% in Temple Lake to 66.6% in Olive Lake), while the control lakes fluctuated through time, or stayed relatively constant (Figure 1.3). The linear mixed model showed that both time, and the interaction between treatment and time, were significant ($p \leq 0.012$ and $p \leq 0.008$, respectively), with the harvest lakes experiencing a significant decrease in $\widehat{N}_c$ from 2017 to 2019 ($p \leq 0.004$), and control lakes experiencing no change ($p > 0.05$).

### *Microsatellite Variation*

Of the 1248 individuals genotyped, 9 were removed because of missing data at 5 or more loci, and an additional 9 were removed due to evidence of sample contamination in the allele depth vs size histogram plots. One locus (ssa.27.d19) was monomorphic and was therefore excluded. Two other loci (ssa.06.8 and ssa.04.d5) were removed from subsequent analyses due to, respectively, a high percentage of ungenotyped individuals (18.8%) and a difficulty in scoring alleles at large allele sizes. This resulted in n=1230 individuals.

For all within-sample microsatellite analyses, each population-year replicate was treated as a unique "population" (i.e., Helen Lake 2017 was treated as a separate population from Helen

Lake 2019 and Helen Lake 2020). Microchecker detected null alleles at locus ssa.20.d16 in a majority of the population-year replicates (13 of 18), therefore we removed this locus from subsequent analyses. Four additional loci had null alleles detected in 2 to 4 of the 18 population-year replicates, but these were spread across lakes and years, with no obvious trends. Linkage disequilibrium was detected between locus ssa-1.7 and two other loci in 7 of the 18 population-year replicates after Bonferroni correction, therefore locus ssa-1.7 was removed from subsequent analyses. Only two other pairs of loci had significant linkage in more than one population-year replicate, but both pairs were only significant in two of the 18 replicates, so no additional loci were removed. Only 25 out of 504 tests for heterozygote excess (<5%), and only 34 of 504 tests for heterozygote deficiency (<7%) were significant at the $p < 0.05$ level. The final dataset therefore consisted of n = 1230 individuals genotyped at 28 microsatellite loci with an amplification success ≥ of 99.7%. Across all lakes and years, the average expected heterozygosity was 0.422 (0.36 – 0.479), average observed heterozygosity was 0.426 (0.364 – 0.485), and mean number of alleles per locus was 3.061 (2.364 – 3.429) (Table S4).

*Changes in $\widehat{N}_b$ and $\widehat{N}_b/\widehat{N}_c$*

The estimates of $\widehat{N}_b$ ranged from 21.5 (McNair Lake 2020) to 115.6 (Helen Lake 2019) (Table S5). All estimates had finite confidence intervals, indicating that the sample size used was sufficient in generating an accurate $N_b$ estimate. The linear mixed model with lake as a random effect showed no significant effect of treatment, time, or their interaction (all $p > 0.05$), indicating that $\widehat{N}_b$ did not change in a predictable way in the harvest lakes compared to the control lakes. Although we did not find statistical significance in the linear mixed model, we can examine the temporal pattern in $\widehat{N}_b$ in each lake using the 95% CIs (Figure 1.4). Mud Lake was

the only lake with non-overlapping CIs through time (indicating it underwent a significant change in $\widehat{N}_b$). Mud experienced a decrease in $\widehat{N}_b$ from 2017 to 2019, and then $\widehat{N}_b$ remained low in 2020 (no change between 2019 and 2020). All other lakes had overlapping CIs for all three years of sampling (Figure 1.4).

Estimates of the $\widehat{N}_b/\widehat{N}_c$ ratio were all well below 1, ranging from 0.014 (Temple 2017) to 0.181 (Helen 2019) (Table S6). On average, the harvest lakes experienced a ten-fold increase in $\widehat{N}_b/\widehat{N}_c$ from 2017 to 2020 (ranging from four- to sixteen-fold), while there was no consistent trend in the control lakes (Figure 1.5). The linear mixed model showed a significant effect of both time ($p \leq 0.046$) and the interaction between treatment and time ($p \leq 0.03$). The standardized residuals approximated a normal distribution, although they were not heteroskedastic, indicating there are likely other predictor variables influencing $\widehat{N}_b/\widehat{N}_c$ that we did not include in this model. In post-hoc analyses of time, we found a significant increase in $\widehat{N}_b/\widehat{N}_c$ between 2017 and 2020 ($p \leq 0.039$), but this was driven solely by the harvest lakes. The post-hoc analysis of the interaction term showed that $\widehat{N}_b/\widehat{N}_c$ in harvest lakes increased between 2017 and 2020 ($p \leq 0.017$), but there was no change in the control lakes ($p > 0.05$). Therefore, consistent with our predictions of genetic compensation, we found increased $\widehat{N}_b/\widehat{N}_c$ through time in the harvest lakes only.

*Influence of Life History on $N_b$*

None of the three models created had any significant terms (i.e., Treatment, $N_c$, CV body length, sex ratio, and any interactions all had $p > 0.05$), and when comparing their AICc, there were no differences between the three models in their ability to predict $N_b$ (the ΔAIC was <1 in all cases; Table 1.1). The backward model selection showed that none of the predictor variables

included in the models were important, and in all cases $\widehat{N}_b \sim 1 + (1 \mid \text{Lake})$ was always the most parsimonious model (see Table S7). Therefore, we did not find any significant influence of life history traits or $N_c$ on $N_b$ in these populations.

**Discussion**

Evaluating the impact of fisheries on both genetic and demographic variables in exploited populations is important in conservation and management, because they can be linked to one another (Waples et al. 2013; Bernos and Fraser 2016; Yates et al. 2017). We experimentally harvested wild brook trout populations from independent lakes to assess how size-selective harvest over three consecutive years influences $N_b$ and $N_c$, and whether harvest-induced changes in demographic or life history traits resulted in a change in $N_b$. Consistent with our expectations, we found that the census size rapidly decreased in all harvest lakes, and we found evidence of genetic compensation in harvested populations, where change in $\widehat{N}_b$ is buffered by density-dependent processes, and $\widehat{N}_b/\widehat{N}_c$ increases as population size decreases. However, contrary to our predictions, we found no link between harvest-induced changes in demographic and life history traits and $\widehat{N}_b$. These results suggest that while demographic variables (like census size) changed rapidly after harvest, the effective size was resilient to short-term change likely due to density-dependent processes such as genetic compensation.

*Demographic responses to harvest*

Our data show strong support for the hypothesis that census size should decrease after three years of successive harvesting. It is well-known that fishing pressure can lead to reductions in census size, as guidelines have been created for commercial fisheries associated with a

maximum sustainable yield, or a maximum reduction in population size relative to their natural abundance (Froese et al. 2016). Our treatment lakes experienced a rapid decline of 60.8% over three years while experiencing an average annual harvest rate of ~0.59. In natural systems exposed to commercial harvest, greater declines have been seen with similar or even lower harvest rates, but the change took place over a greater timespan. In the Baltic Sea, Baltic cod stocks experienced an ~80% decline over a ten-year period (1981 to 1992) while the estimated mean harvest rate was 0.56 (Jonzén et al. 2001). Similarly, red grouper in the Campeche Bank declined 88% from 1986 to 2001, but fishing mortality only reached an average of 0.37 at the end of the 15-year period (Giménez-Hurtado et al. 2005). While there is less data on recreational fishing, Cahill et al. (2021) showed mean fishing mortality rates of ~0.3 in recreational Walleye fisheries in Alberta, Canada, from 1980 to 1996, which contributed to a collapse in these populations (a "collapse" is generally considered an 80% or more decline in population size; Post et al., 2002). Therefore, while the decline in $\widehat{N}_c$ in our system was markedly more rapid, it is similar to those seen in commercially and recreationally harvested populations.

*Harvest-induced changes in effective size and evidence of genetic compensation*

We found no consistent change in $\widehat{N}_b$ across control or harvest lakes, and our model showed no statistical significance between control and harvest lakes. However, we did find that one harvest lake, Mud, experienced a significant decrease in $\widehat{N}_b$. These results actually agree with our initial hypothesis, as we expected that in harvest lakes, change in $\widehat{N}_b$ would be buffered by density-dependent processes, resulting in no change in $\widehat{N}_b$ (i.e., in Olive and Temple) in at least two of the three lakes. The buffering of change in $\widehat{N}_b$ is much more obvious when assessing the change in $\widehat{N}_b/\widehat{N}_c$; we found that harvest lakes all increased in $\widehat{N}_b/\widehat{N}_c$, indicating that genetic

compensation has occurred. Although the change in $\widehat{N}_b/\widehat{N}_c$ is driven by decreases in $\widehat{N}_c$ (rather than increases in $\widehat{N}_b$), this is still evidence of genetic compensation, as the proportional reductions in $\widehat{N}_b$ are much smaller than expected based on a reduced $\widehat{N}_c$. This relationship also holds true if we convert our $\widehat{N}_b$ estimates to $\widehat{N}_e$ using life history trait conversions from Waples et al. (2014). With converted $\widehat{N}_e$, there were still no consistent changes in effective size, but the ratios of $\widehat{N}_e/\widehat{N}_c$ in harvest lakes all increased (see Appendix B for more information). Similar results were found by Kuparinen et al. (2016) in a modelling study on Atlantic cod, where they showed that increasing adult mortality reduces both $N_c$ and $N_e$, however the ratio between $N_e$ and $N_c$ increases.

Genetic compensation has been described in small populations from several different species, including several salmonid populations, and is generally attributed to a lower variance in breeding success at smaller population sizes (Ardren and Kapuscinski 2003; Jehle et al. 2005; Watts et al. 2007; Beebee 2009; Saarinen et al. 2010; Bernos and Fraser 2016). In salmonid populations, females often compete over high-quality breeding sites, resulting in nest destruction or superimposition (Fleming and Reynolds 2004). Therefore, at larger population sizes, variance in reproductive success should be larger due to competition for breeding sites (Chebanov 1991). In unexploited, stream-dwelling brook trout populations, Bernos and Fraser (2016) found a relationship between stream length, $N_b$, and $N_c$, indicating that breeding habitat availability is likely a factor in density-dependent processes in brook trout. While our populations in this study are lake-dwelling, many of them spawn in inlets or outlets, therefore the relationship between spawning habitat availability and $N_b/N_c$ is likely similar.

In addition to the competition for breeding sites, density-dependence may also have an impact on life history traits that could result in a high $N_b/N_c$. At lower densities, individuals may

have higher growth, earlier maturation, or increased fecundity (Rose et al. 2001; Johnston and Post 2009; Matte et al. 2020), all of which could allow for the harvested populations to maintain a high $N_b/N_c$. Density-dependent effects on $N_b$ might also be population-specific. We found differences between the three harvest lakes in their response (i.e., Mud decreased in $\widehat{N}_b$ while Olive and Temple did not change), which might be attributable to different density-dependent responses. Matte et al. (2020) showed that in brook trout populations, density-dependent responses in growth and mortality varied, even after accounting for environmental differences. Although we cannot make conclusions about whether these underlying mechanisms influenced our populations, density-dependent genetic compensation likely played a role in buffering change in $N_b$ in the harvest lakes.

*Life history traits and other influences on $N_b$*

Contrary to our prediction, we did not find any evidence that the sex ratio or CV of body size had an influence on $\widehat{N}_b$ in these populations. Our predictions were based on a priori knowledge that an uneven sex ratio usually leads to a reduction in effective size (Franklin 1980), and that populations with a wider range in body size may benefit from competitive release if the largest individuals are removed. However, we did not consider whether other life-history traits might be more important in predicting $N_e$, especially in the context of exploited populations. Waples et al. (2013) showed that age at maturity and adult lifespan explain up to half of the variance in $N_e/N_c$ across 63 iteroparous species (including two salmonid species). In the context of size-selective harvest, age at maturity and adult lifespan would also make sense as potential drivers of change in $N_b$, as these two life-history traits will likely be impacted by harvest (Law 2000; Enberg et al. 2012). Indeed, other work from our project found that as density decreased,

both males and females had an earlier maturation, especially those from age 0+, 1+, and 2+ age-classes (Matte et al., submitted). Therefore, age-at-maturity may have helped to predict changes in effective size, however it may not have resulted in significant results, simply due to the lack of change in $\widehat{N}_b$ across control and harvest lakes in our study. Matte et al. (submitted) also looked at changes in size-frequency distribution of body length. They found that two of the harvest lakes in this study, Temple and Olive, were stunted populations with low mean and variance in body size and did not experience much change in body size variance, whereas Mud decreased in variance and body sizes became smaller. Although in our analysis we did not find a statistically significant influence of variance in body size (perhaps due to sample size limitations), the results from Matte et al. (submitted) may help explain why Mud was the only harvest lake in this study to experience a decrease in $\widehat{N}_b$.

*Caveats and future research directions*

Studies like ours that experimentally harvest wild populations with proper controls are exceedingly rare and sorely needed to inform fisheries management (Kuparinen and Merilä 2007; Audzijonyte et al. 2013; Heino et al. 2015). Nevertheless, our work has a few caveats that are important to take into account when interpreting the results. Firstly, despite three- or four-month long field seasons with 4+ crew members, the replication in this study was relatively low. Three years of data from six lakes only yielded 18 data points for the $\widehat{N}_b$ and $\widehat{N}_c$ datasets, therefore we had low power to test for significance in our models. This could explain why we did not find any statistically significant effect of harvest on $\widehat{N}_b$ or find a link between life history traits and $\widehat{N}_b$, however in this discussion we are careful to limit our conclusions to only what the data show us. Related to this replication issue is that our work was conducted over a relatively

short temporal scale. There could be a time-lag in detecting changes in $N_b$, especially if density-dependent processes only buffer change in the short-term. If $N_c$ continued to decrease in these populations, it is possible they could experience inbreeding, and subsequently a decrease in effective size (Wang et al. 2002). Finally, our results may be a more general consequence of increased mortality, rather than from size-selectivity of fishing (Garcia et al. 2012), as the genetic compensation likely resulted from a change in density. Future work would benefit from comparing balanced harvest (i.e., harvesting across all size classes) to size-selective harvest to compare impacts on effective size, and ideally increase replication over both time and space to increase statistical power.

*Conclusions*

Assessing the impact of fisheries on the effective size of exploited populations can reveal important impacts on both demographic and genetic processes. Here, we show how in populations of size-selectively harvested brook trout, that demographic parameters changed rapidly, while the effective size experienced little or no change. Over the short-term, change in effective size is likely buffered by density-dependent processes, but this does not mean that the populations are immune to decreases in $N_b$ (as seen in Mud Lake). Even though $N_b/N_c$ may increase in a harvested population, it could still experience a decrease in effective size, increasing the risk of inbreeding, and reducing the fitness of the population (Charlesworth and Willis 2009). The risk of inbreeding in harvested populations would become especially prevalent over the long-term if population size were to decline further. To our knowledge, this study is also the first to show the effects of harvest in natural populations with relatively small effective and census sizes. We showed similar results to the modelling study by Kuparinen et al. (2016),

however our effective sizes were much lower (ranging from $\widehat{N}_b$ of ~20 to 65 in harvest lakes which would translate into $N_e$ of <100 to a few hundred, compared to $N_e$ of ~500 to 2000 in the modelling study), and our populations were in a lacustrine rather than a marine environment. In the context of conservation and management, showing these relationships at small population sizes is especially important, as small populations are often highly managed, and may be at a high risk of extirpation (Lande 1993; Meuwissen 2009). In sum, although effective size of harvested populations may be resilient to change in the short-term, proper monitoring and management of fisheries is important to ensure the sustainability of fisheries, especially in the long-term.

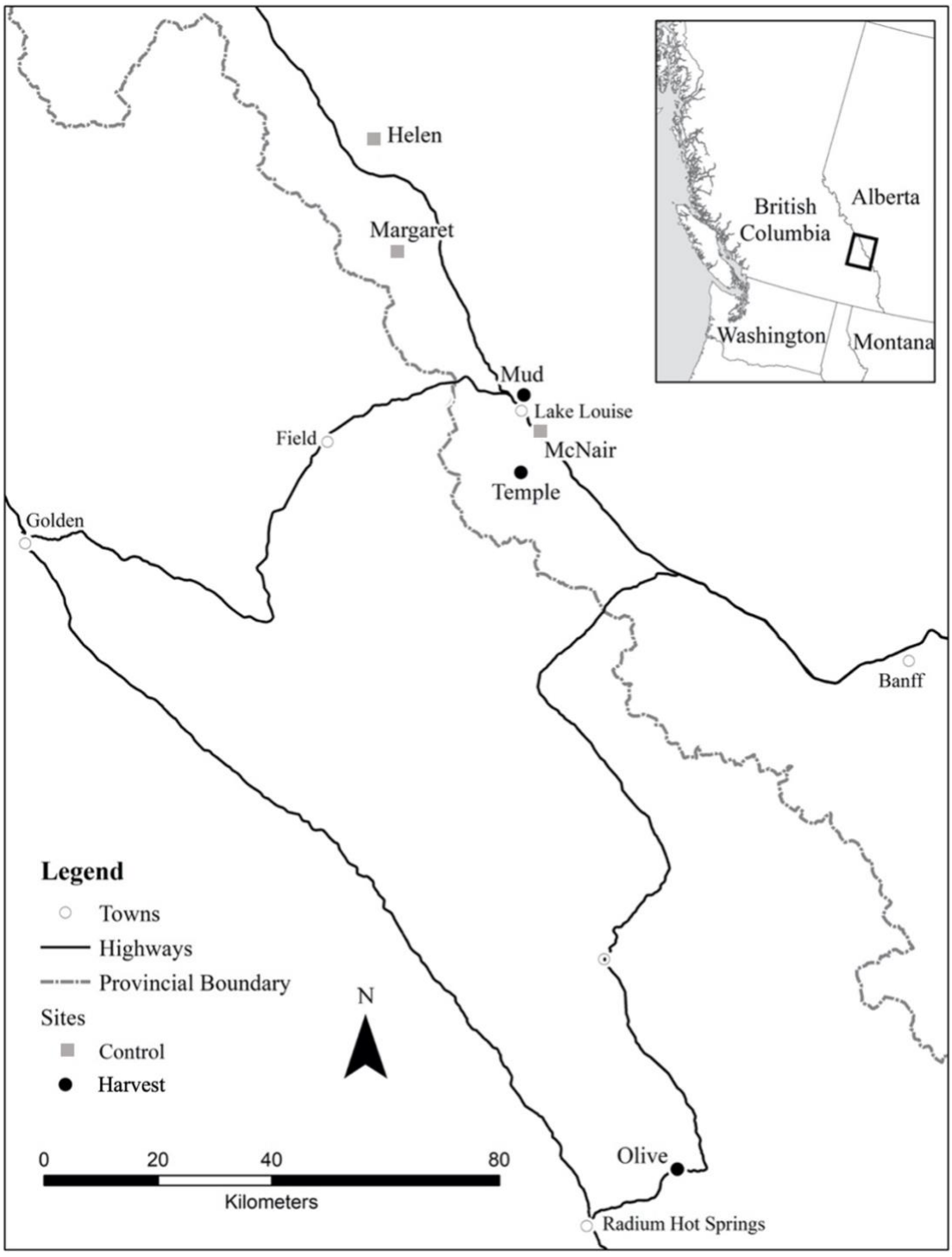


**Figure 1.1:** Map of six brook trout populations in alpine lakes in Banff and Kootenay National Parks in the Canadian Rocky Mountains

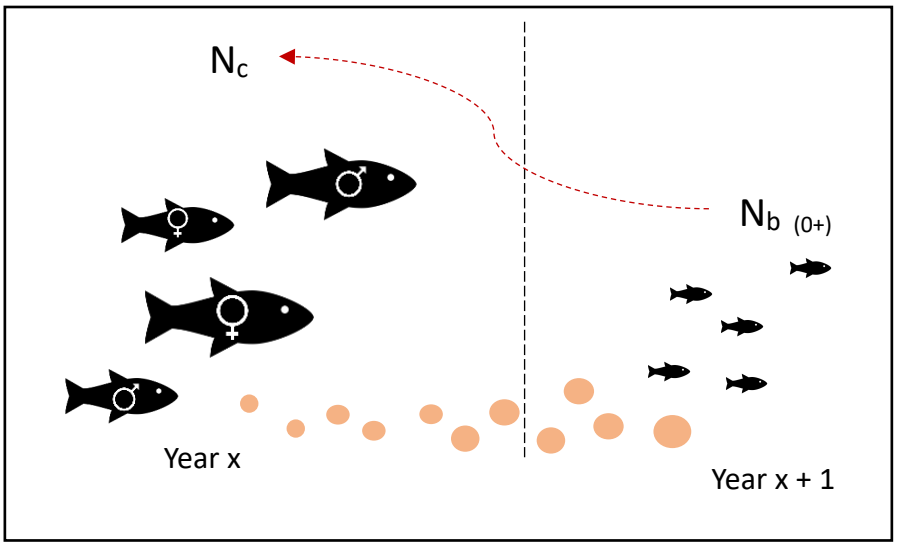


**Figure 1.2**: Correct linking of effective number of breeders ($N_b$) to corresponding census size ($N_c$) in Brook Trout, where spawning occurs in the fall and fry emerge in the following spring (with the orange dots representing the eggs). $N_b$ of a 0+ (young-of-the-year) cohort should be linked with the $N_c$ of the adults in the previous year that produced the cohort.

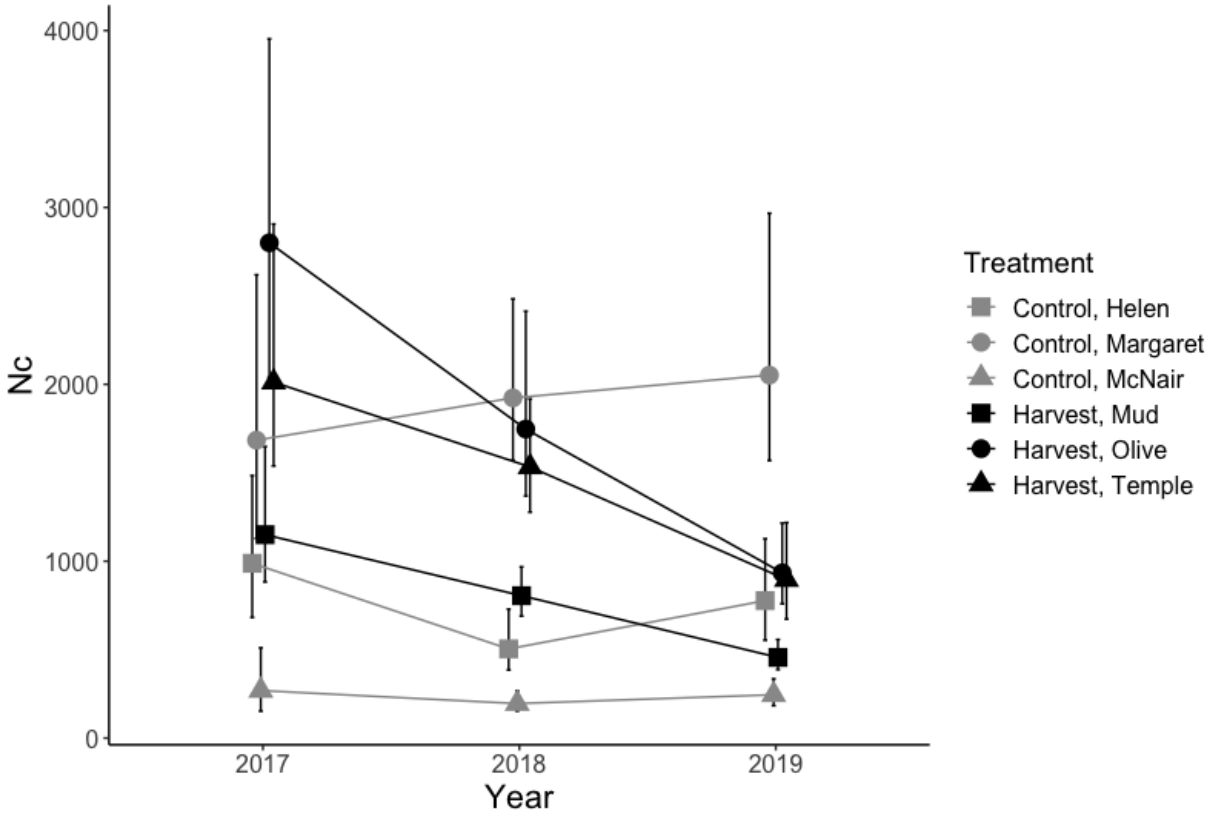


**Figure 1.3**: Census size ($\widehat{N}_c$) estimates for control and harvest brook trout populations in the Rocky Mountains from 2017 to 2019. Error bars indicate 95% confidence intervals

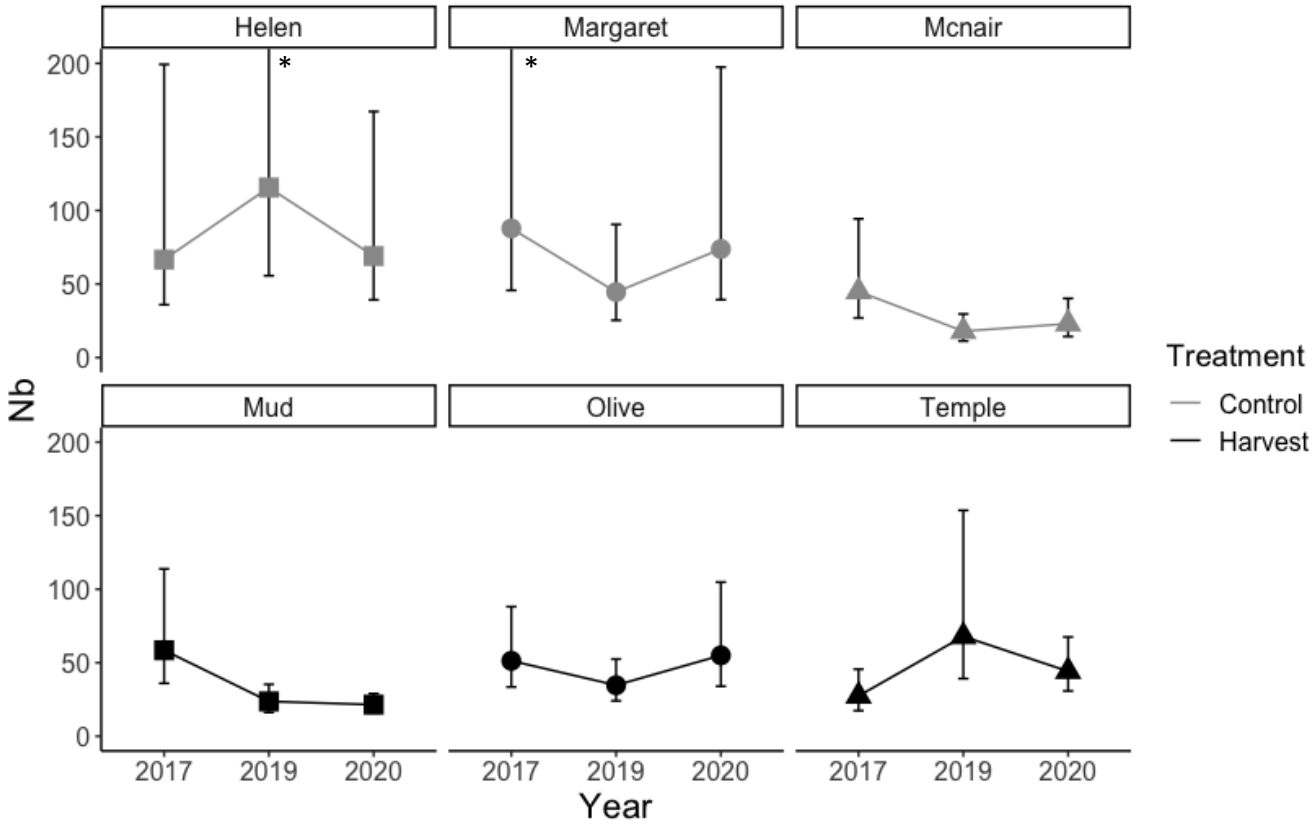


**Figure 1.4**: Effective number of breeders ($\widehat{N}_b$) change over time in control and harvest populations of brook trout in the Rocky Mountains. Error bars indicate 95% confidence intervals calculated using the jackknife method. Asterisks indicate where the error bar did not fit within the limits of the graph.

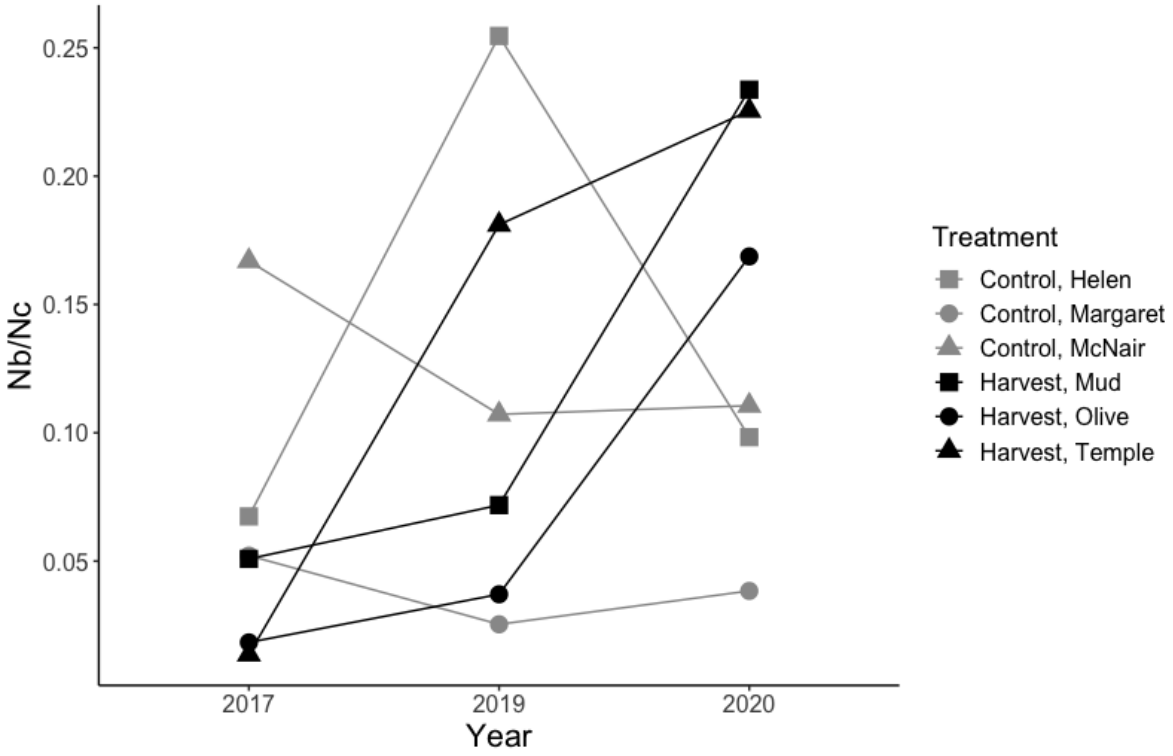


**Figure 1.5**: Effective to census size ratios ($\widehat{N}_b/\widehat{N}_c$) over time for control and harvest brook trout populations in the Rocky Mountains.

**Table 1.1**: AIC values of Generalized Linear Mixed Models comparing the influence of census size ($N_c$) and life history traits on effective number of breeders ($N_b$) in control and harvest alpine brook trout populations

| Model Terms | Log Likelihood | AIC | ΔAIC |
|---|---|---|---|
| $N_b$ ~ CV Body Length * Treatment + (1 \| Lake) | – 79.96 | 179.56 | 0.00 |
| $N_b$ ~ Sex Ratio * Treatment + (1 \| Lake) | – 80.05 | 179.74 | 0.18 |
| $N_b$ ~ $N_c$ * Treatment + (1 \| Lake) | – 80.41 | 180.46 | 0.90 |

**References**

Adams, S. B., C. A. Frissell, and B. E. Rieman. 2000. Movements of nonnative brook trout in relation to stream channel slope. Transactions of the American Fisheries Society 129(3):623–638. Taylor & Francis.

Allen, M. S., C. J. Walters, and R. Myers. 2008. Temporal Trends in Largemouth Bass Mortality, with Fishery Implications. North American Journal of Fisheries Management 28(2):418–427.

Allendorf, F., P. England, G. Luikart, P. Ritchie, and N. Ryman. 2008. Genetic effects of harvest on wild animal populations. Trends in Ecology & Evolution 23(6):327–337.

Allendorf, F. W., and J. J. Hard. 2009. Human-induced evolution caused by unnatural selection through harvest of wild animals. Proceedings of the National Academy of Sciences 106:9987–9994.

Anderson, J. H., P. L. Faulds, W. I. Atlas, G. R. Pess, and T. P. Quinn. 2010. Selection on breeding date and body size in colonizing coho salmon, Oncorhynchus kisutch: selection in colonizing coho salmon. Molecular Ecology 19:2562–2573.

Ardren, W. R., and A. R. Kapuscinski. 2003. Demographic and genetic estimates of effective population size (Ne) reveals genetic compensation in steelhead trout: steelhead trout effective population size. Molecular Ecology 12(1):35–49.

Audzijonyte, A., A. Kuparinen, and E. A. Fulton. 2013. How fast is fisheries-induced evolution? Quantitative analysis of modelling and empirical studies. Evolutionary Applications 6(4):585–595.

Bates, D. 2011. Mixed models in R using the lme4 package Part 6: Nonlinear mixed models. Statistics:1–9.

Bates, D., M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, B. Dai, F. Scheipl, G. Grothendieck, P. Green, J. Fox, A. Bauer, and P. N. Krivitsky. 2020. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-26.

Beebee, T. J. C. 2009. A comparison of single-sample effective size estimators using empirical toad ( *Bufo calamita* ) population data: genetic compensation and population size-genetic diversity correlations. Molecular Ecology 18(23):4790–4797.

Bernos, T. A., and D. J. Fraser. 2016. Spatiotemporal relationship between adult census size and genetic population size across a wide population size gradient. Molecular Ecology 25(18):4472–4487.

Bernos, T. A., M. C. Yates, and D. J. Fraser. 2018. Fine-scale differences in genetic and census population size ratios between two stream fishes. Conservation Genetics 19(2):265–274.

Blanchfield, P. J., and M. S. Ridgway. 2005. The relative influence of breeding competition and habitat quality on female reproductive success in lacustrine brook trout (Salvelinus fontinalis). Canadian Journal of Fisheries and Aquatic Sciences 62:2694–2705.

Blanchfield, P. J., M. S. Ridgway, and C. C. Wilson. 2003. Breeding success of male brook trout (Salvelinus fontinalis) in the wild: paternity in lacustrine brook trout. Molecular Ecology 12(9):2417–2428.

Brookes, B., H.B. Jeon, A.M. Derry, J.R. Post, S.M. Rogers, S. Humphries, and D.J. Fraser. 2022. Neutral and adaptive drivers of genomic change in introduced brook trout (Salvelinus fontinalis) populations revealed by pooled sequencing. Ecology and evolution, 12(2), p.e8584.

Cahill, C.L., C.J. Walters, A.J. Paul, M.G. Sullivan, and J.R. Post. 2021. Unveiling the recovery dynamics of walleye after the invisible collapse. Canadian Journal of Fisheries and Aquatic Sciences 00: 1–16 (0000). dx.doi.org/10.1139/cjfas-2021-0065

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nature Reviews Genetics 10(11):783–796.

Chebanov, N. A. 1991. The effect of spawner density on spawning success, egg survival, and size structure of the progeny of the sockeye salmon, Oncorhynchus nerka. Journal of Ichthyology 31:101–106.

Dare, M. R. 2003. Mortality and long-term retention of passive integrated transponder tags by spring Chinook salmon. North American Journal of Fisheries Management 23(3):1015–1019. Taylor & Francis.

Diaz, M., D. Wethey, J. Bulak, and B. Ely. 2000. Effect of Harvest and Effective Population Size on Genetic Diversity in a Striped Bass Population. Transactions of the American Fisheries Society 129(6):1367–1372.

Do, C., R. S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources 14(1):209–214.

Enberg, K., C. Jørgensen, E. S. Dunlop, Ø. Varpe, D. S. Boukal, L. Baulier, S. Eliassen, and M. Heino. 2012. Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence: Fishing-induced evolution of growth. Marine Ecology 33(1):1–25.

Excoffier, L., and H. E. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources 10(3):564–567. Wiley Online Library.

Fenberg, P. B., and K. Roy. 2008. Ecological and evolutionary consequences of size-selective harvesting: how much do we know? Molecular Ecology 17(1):209–220.

Ferchaud, A.-L., C. Perrier, J. April, C. Hernandez, M. Dionne, and L. Bernatchez. 2016. Making sense of the relationships between Ne, Nb and Nc towards defining conservation thresholds in Atlantic salmon (Salmo salar). Heredity 117(4):268–278.

Fleming, I. A., and J. D. Reynolds. 2004. Salmonid breeding systems. Evolution illuminated: salmon and their relatives:264–294. Oxford University Press Oxford.

Franklin, I. R. 1980. Evolutionary change in small populations. Pages 135–149 *in* M. E. Soule and B. A. Wilcox, editors. Conservation Biology - An Evolutionary-Ecological Perspective.

Fraser, D. J. 2013. The emerging synthesis of evolution with ecology in fisheries science. Canadian Journal of Fisheries and Aquatic Sciences 70(9):1417–1428.

Froese, R., H. Winker, D. Gascuel, U. R. Sumaila, and D. Pauly. 2016. Minimizing the impact of fishing. Fish and Fisheries 17(3):785–802.

Garcia, S. M., J. Kolding, J. Rice, M.-J. Rochet, S. Zhou, T. Arimoto, J. E. Beyer, L. Borges, A. Bundy, D. Dunn, E. A. Fulton, M. Hall, M. Heino, R. Law, M. Makino, A. D. Rijnsdorp, F. Simard, and A. D. M. Smith. 2012. Reconsidering the Consequences of Selective Fisheries. Science 335(6072):1045–1047.

Giménez-Hurtado, E., R. Coyula-Pérez-Puelles, S. E. Lluch-Cota, A. A. González-Yañez, V. Moreno-García, and R. Burgos-de-la-Rosa. 2005. Historical biomass, fishing mortality, and recruitment trends of the Campeche Bank red grouper (Epinephelus morio). Fisheries Research 71(3):267–277.

Gobin, J., N. P. Lester, M. G. Fox, and E. S. Dunlop. 2016. Effects of changes in density-dependent growth and recruitment on sustainable harvest of lake whitefish. Journal of Great Lakes Research 42(4):871–882.

Gray, Q. Z., D. J. Fraser, and J. W. A. Grant. 2014. Extirpation for Conservation: Applying Predictors of Extinction Risk to Eradicate Introduced Trout Populations for Lake Restoration. Ecological Restoration 32(1):59–67.

Gries, G., and B. H. Letcher. 2002. Tag retention and survival of age-0 Atlantic salmon following surgical implantation with passive integrated transponder tags. North American Journal of Fisheries Management 22(1):219–222. Taylor & Francis.

Heino, M., B. Díaz Pauli, and U. Dieckmann. 2015. Fisheries-Induced Evolution. Annual Review of Ecology, Evolution, and Systematics 46(1):461–480.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical Research 38(3):209–216.

Huggins, R., and A. Chao. 2005. Classical Closed-population Capture-Recapture Models. Pages 22–35 Handbook of Capture-Recapture Analysis. Princeton University Press.

Jehle, R., G. A. Wilson, J. W. Arntzen, and T. Burke. 2005. Contemporary gene flow and the spatio-temporal genetic structure of subdivided newt populations (Triturus cristatus, T. marmoratus): Gene flow and drift in newts. Journal of Evolutionary Biology 18(3):619–628.

Johnston, F. D., and J. R. Post. 2009. Density-dependent life-history compensation of an iteroparous salmonid. Ecological Applications 19(2):449–467.

Jones, M. W., and J. A. Hutchings. 2001. The influence of male parr body size and mate competition on fertilization success and effective population size in Atlantic salmon. Heredity 86:675–684.

Jonzén, N., P. Lundberg, M. Cardinale, and F. Arrhenius. 2001. Variable fishing mortality and the possible commercial extinction of the eastern Baltic cod. Marine Ecology Progress Series 210:291–296.

Kendall, N. W., and T. P. Quinn. 2013. Size-selective fishing affects sex ratios and the opportunity for sexual selection in Alaskan sockeye salmon *Oncorhynchus nerka*. Oikos 122(3):411–420.

Kuparinen, A., and M. Festa-Bianchet. 2017. Harvest-induced evolution: insights from aquatic and terrestrial systems. Philosophical Transactions of the Royal Society B: Biological Sciences 372(1712):20160036.

Kuparinen, A., J. A. Hutchings, and R. S. Waples. 2016. Harvest-induced evolution and effective population size. Evolutionary Applications 9(5):658–672.

Kuparinen, A., D. M. Keith, and J. A. Hutchings. 2014. Allee Effect and the Uncertainty of Population Recovery: Allee Effect and Population Recovery. Conservation Biology 28(3):790–798.

Kuparinen, A., and J. Merilä. 2007a. Detecting and managing fisheries-induced evolution. Trends in Ecology & Evolution 22(12):652–659.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. The American Naturalist 142(6):911–927. University of Chicago Press.

Law, R. 2000. Fishing, selection, and phenotypic evolution. ICES Journal of Marine Science 57(3):659–668.

Lee, A. M., S. Engen, and B.-E. Saether. 2011. The influence of persistent individual differences and age at maturity on effective population size. Proceedings of the Royal Society B: Biological Sciences 278(1722):3303–3312.

Lenth, R. 2018. Package ‘lsmeans’. R package version 2.3.0-0.

Letcher, B. H., K. H. Nislow, J. A. Coombs, M. J. O'Donnell, and T. L. Dubreuil. 2007. Population Response to Habitat Fragmentation in a Stream-Dwelling Brook Trout Population. PLoS ONE 2(11):e1139.

Matte, J., D. J. Fraser, and J. W. A. Grant. 2020. Population variation in density-dependent growth, mortality and their trade-off in a stream fish. Journal of Animal Ecology 89(2):541–552.

Meuwissen, T. 2009. Genetic management of small populations: A review. Acta Agriculturae Scand Section A 59(2):71–79. Taylor & Francis.

Ogle, D., P. Wheeler, and A. Dinno. 2019. FSA: Fisheries Stock Analysis.

Ostrand, K. G., G. B. Zydlewski, W. L. Gale, and J. D. Zydlewski. 2011. Long term retention, survival, growth, and physiological indicators of juvenile salmonids marked with passive integrated transponder tags. Page American Fisheries Society Symposium.

Ovenden, J. R., O. Berry, D. J. Welch, R. C. Buckworth, and C. M. Dichmont. 2015. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. Fish and Fisheries 16(1):125–159. Wiley Online Library.

Palstra, F. P., and D. J. Fraser. 2012. Effective/census population size ratio estimation: a compendium and appraisal. Ecology and Evolution 2(9):2357–2365.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, S. Heisterkamp, B. Van Willigen, and R. Maintainer. 2017. Package 'nlme.' Linear and nonlinear mixed effects models, version 3(1).

Post, J.R., M. Sullivan, S. Cox, N.P. Lester, C.J. Walters, E.A. Parkinson, A.J. Paul, L. Jackson, and B.J. Shuter. 2002. Canada's recreational fisheries: the invisible collapse?. Fisheries, 27(1), pp.6-17.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rose, K. A., J. H. Cowan, K. O. Winemiller, R. A. Myers, and R. Hilborn. 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis: Compensation in fish populations. Fish and Fisheries 2(4):293–327.

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular ecology resources 8(1):103–106. Wiley Online Library.

Ruzzante, D. E., G. R. McCracken, B. Førland, J. MacMillan, D. Notte, C. Buhariwalla, J. Mills Flemming, and H. Skaug. 2019. Validation of close-kin mark–recapture (CKMR) methods for estimating population abundance. Methods in Ecology and Evolution 10(9):1445–1453.

Ruzzante, D. E., G. R. McCracken, S. Parmelee, K. Hill, A. Corrigan, J. MacMillan, and S. J. Walde. 2016. Effective number of breeders, effective population size and their relationship with census size in an iteroparous species, *Salvelinus fontinalis*. Proceedings of the Royal Society B: Biological Sciences 283(1823):20152601.

Saarinen, E. V., J. D. Austin, and J. C. Daniels. 2010. Genetic estimates of contemporary effective population size in an endangered butterfly indicate a possible role for genetic compensation: Effective population size estimators. Evolutionary Applications 3(1):28–39.

Shrimpton, J. M., and D. D. Heath. 2003. Census vs. effective population size in chinook salmon: large- and small-scale environmental perturbation effects. Molecular Ecology 12(10):2571–2583.

Stegmann, G., R. Jacobucci, J. R. Harring, and K. J. Grimm. 2018. Nonlinear mixed-effects modeling programs in R. Structural Equation Modeling: A Multidisciplinary Journal 25(1):160–165. Taylor & Francis.

Stokes, K., and R. Law. 2000. Fishing as an evolutionary force. Marine Ecology Progress Series 208:307–309.

Thomas, G., H. Quoß, J. Hartmann, and R. Eckmann. 2009. Human-induced changes in the reproductive traits of Lake Constance common whitefish (Coregonus lavaretus). Journal of Evolutionary Biology 22(1):88–96. Wiley Online Library.

Thompson, P. D., and F. J. Rahel. 1998. Evaluation of artificial barriers in small Rocky Mountain streams for preventing the upstream movement of brook trout. North American Journal of Fisheries Management 18(1):206–210. Taylor & Francis.

Topping, D. T., and S. T. Szedlmayer. 2013. Use of ultrasonic telemetry to estimate natural and fishing mortality of red snapper. Transactions of the American Fisheries Society 142(4):1090–1100. Taylor & Francis.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4(3):535–538.

Wang, S., J. J. Hard, and F. Utter. 2002. Salmonid inbreeding: a review. Reviews in Fish Biology and Fisheries 11(4):301–319. Springer.

Waples, R. S. 2002. Definition and estimation of effective population size in the conservation of endangered species. Population viability analysis:147–168. University of Chicago Press.

Waples, R. S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. Conservation Genetics 7(2):167–184.

Waples, R. S., and T. Antao. 2014. Intermittent breeding and constraints on litter size: consequences for effective population size per generation ( $N_e$ ) and per reproductive cycle ( $N_b$ ): life history variants and $N_e$. Evolution 68(6):1722–1734.

Waples, R. S., T. Antao, and G. Luikart. 2014. Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. Genetics 197(2):769–780.

Waples, R. S., and C. Do. 2008. LDNE : a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8(4):753–756.

Waples, R. S., G. Luikart, J. R. Faulkner, and D. A. Tallmon. 2013. Simple life-history traits explain key effective population size ratios across diverse taxa. Proceedings of the Royal Society B: Biological Sciences 280(1768):20131339.

Watts, P. C., I. J. Saccheri, S. J. Kemp, and D. J. Thompson. 2007. Effective population sizes and migration rates in fragmented populations of an endangered insect (Coenagrion mercuriale: Odonata). Journal of Animal Ecology 76(4):790–800.

Whiteley, A. R., J. A. Coombs, M. Cembrola, M. J. O'Donnell, M. Hudy, K. H. Nislow, and B. H. Letcher. 2015. Effective number of breeders provides a link between interannual variation in stream flow and individual reproductive contribution in a stream salmonid. Molecular Ecology 24(14):3585–3602.

Yates, M. C., T. A. Bernos, and D. J. Fraser. 2017. A critical assessment of estimating census population size from genetic population size (or vice versa) in three fishes. Evolutionary Applications 10(9):935–945.

Yue, G. H., and L. Orban. 2005. A simple and affordable method for high-throughput DNA extraction from animal tissues for polymerase chain reaction. ELECTROPHORESIS 26(16):3081–3083.

Zhan, L., I. G. Paterson, B. A. Fraser, B. Watson, I. R. Bradbury, P. Nadukkalam Ravindran, D. Reznick, R. G. Beiko, and P. Bentzen. 2017. MEGASAT: automated inference of microsatellite genotypes from sequence data. Molecular Ecology Resources 17(2):247–256.

Zipkin, E. F., P. J. Sullivan, E. G. Cooch, C. E. Kraft, B. J. Shuter, and B. C. Weidel. 2008. Overcompensatory response of a smallmouth bass (Micropterus dolomieu) population to harvest: release from competition? Canadian Journal of Fisheries and Aquatic Sciences 65(10):2279–2292.

## Chapter 2: Global assessment of effective population sizes reveals deficiencies in meeting 50/500 criteria across taxa

Shannon H. Clarke, Elizabeth Lawrence, Jean-Michel Matte, Brian Gallagher, Sarah Salisbury, Sozos Michaelides, Ramela Koumrouyan, Daniel E. Ruzzante, James W.A. Grant, Dylan J. Fraser

## Abstract

Integrating genetic measures into extinction risk assessments is imperative for reducing biodiversity loss and ensuring the viability of wild populations. Effective population size ($N_e$) is particularly useful for conservation, as it predicts the rates of genetic drift, inbreeding, and adaptive potential of a population. Current guidelines recommend minimum $N_e$ of 50 and 500 to avoid short-term inbreeding and long-term adaptive potential, respectively. However, a global assessment of the extent to which wild populations reach these thresholds has not been conducted, nor has the relationship between $N_e$ and human activities been assessed. Here, we show that certain taxonomic groups are less likely to meet the 50/500 thresholds and are disproportionately impacted by human activities. Through a quantitative review, we found consistent taxonomic differences in $N_e$, and that plant, mammal, and amphibian populations had a ~50% or less chance of reaching $N_e = 50$ and a <5% chance of reaching 500. Across all taxonomic groups, $N_e$ was reduced in areas of greater human footprint, especially in amphibians and mammals. Our results suggest that current $N_e$ guidelines may not be representative or appropriate for all taxonomic groups. These findings can help inform changes to listing criteria or prioritize assessment of populations from taxa most at risk of falling below conservation thresholds.

## Introduction

Wild populations are facing increasing threats from human activities (Bellard et al., 2012; Rosser & Mainka, 2002; Wilson et al., 2016). As a result, the number of endangered species has doubled in the last decade (IUCN Red List, 2020), and we are currently in the midst of a sixth mass extinction event (Ceballos et al., 2015; Pimm et al., 2014). Measuring the extinction risk of populations is therefore important to maintain wild populations and reduce biodiversity loss. Extinction risk is commonly evaluated through demographic and ecological measures such as declines in population size or contractions in range size (Mace et al., 2008). However, genetic measures such as genetic diversity or inbreeding depression provide complementary information to demographic or ecological assessments and are therefore important to integrate into evaluation of extinction risk (Dunham et al., 1999; Frankham, 2005).

A key genetic variable for evaluating extinction risk is the effective population size ($N_e$), which relates to the adaptive potential in populations. Not only does $N_e$ influence the amount of genetic diversity lost through random genetic drift within a population, but it also reflects the level of inbreeding, which can be especially prevalent in small, vulnerable populations, and can contribute to extinction risk (Bijlsma et al., 2000; Crow & Kimura, 1970; Waples, 2002a). Indeed, contemporary $N_e$ has been pointed to as a better indicator of extinction risk than neutral genetic diversity of a population, as $N_e$ is more closely related to adaptive potential (Teixeira & Huber, 2021). Thresholds of minimum $N_e$ have been established to ensure wild population persistence. The 50/500 rule recommends a minimum $N_e$ of 50 to avoid inbreeding in the short-term, and a minimum $N_e$ of 500 to allow for adaptation in the long-term, based on genetic principles such as mutation, selection, and inbreeding depression (Franklin, 1980). However, there has been debate over how the 50/500 rule should be applied in conservation, and some

experts recommend increasing the minimum thresholds to 100 and 1000 to maintain populations' fitness and evolutionary potential (see Frankham et al., 2014; García-Dorado, 2015; Jamieson & Allendorf, 2012). Overall, the 50/500 criteria are often incorporated into population viability analyses or minimum viable population sizes for global conservation initiatives (e.g. IUCN Red List, Convention on Biological Diversity; Hoban et al., 2020; Mace et al., 2008).

A global assessment of $N_e$ in wild populations in relation to 50/500 thresholds and human activities is timely due to its potential use in conservation planning and forecasting adaptive responses of at-risk populations to environmental or human stressors. Indeed, $N_e$ estimates have accumulated rapidly in the literature for more than 10 years since previous reviews of $N_e$ were conducted (Frankham, 1995; Palstra & Fraser, 2012; Palstra & Ruzzante, 2008). These previous reviews lacked the sample sizes required to parse out taxonomic trends and were highly biased towards salmonid fishes (e.g. 40% of estimates in Palstra & Ruzzante, 2008). Additionally, these reviews focused heavily on estimates generated using either demographic methods (Frankham, 1995), or temporal genetic methods (Palstra & Fraser, 2012; Palstra & Ruzzante, 2008). Recent developments in the field of $N_e$ estimation allow for the estimation of $N_e$ using a single sample, and the ability to use genome-wide data from SNPs (single-nucleotide polymorphisms) in $N_e$ estimation (Barbato et al., 2015).

With advancements in $N_e$ estimation methods and bias reduction, as well as a rapid increase in the number of estimates, it is timely to revisit $N_e$ in wild populations. Our review summarized single-sample $N_e$ estimates across taxa using systematic review approaches (e.g. O'Dea et al., 2021), and incorporated substantial geo-referencing of populations. Specifically, we: (1) evaluated taxonomic differences in $N_e$ and their likelihood to meet 50/500 thresholds; (2)

created a global map of $N_e$ and assessed the influence of human activities; and (3) compared the ratio between $N_e$ and $N_c$ across taxa.

For objective (1), we generated taxon-specific estimates of $N_e$, as values generated across all taxonomic groups in previous reviews may not be representative for specific populations. $N_e$ is closely linked with demographic characteristics that can vary between taxonomic groups; therefore, we expected that average values would vary between taxa. We also compared $N_e$ estimates from wild populations to the 50/500 rule and compared the likelihood of different taxonomic groups to meet these thresholds. For objective (2) we created global maps of $N_e$ within and across taxa, to assess sampling biases and identify any hotspots. We also tested the hypothesis that $N_e$ is negatively correlated with the Global Human Footprint (WCS & CIESIN, 2005). The Global Human Footprint is a quantitative assessment of human impact and includes the influence of factors such as population density, agriculture, transportation networks, and light pollution (Venter et al., 2016). We expected the footprint to be negatively correlated with $N_e$, as fragmentation, habitat destruction, and pollution can lead to reduced population sizes or impact breeding, therefore reducing $N_e$. Finally, for objective (3), we assessed how the ratio between effective and census size ($N_c$) varied across taxonomic groups. The ratio between $N_e$ and $N_c$ is often used in risk assessments or calculation of minimum viable population sizes, and previous reviews have reported average values ranging from 0.1 to 0.2 (Frankham, 1995; Palstra & Fraser, 2012; Palstra & Ruzzante, 2008). We hypothesized that $N_e/N_c$ would vary between taxonomic groups due to differing life history characteristics, with groups with type III survivorship curves (i.e., high fecundity but low juvenile survival) having the lowest ratios due to a high variance in reproductive success.

## Methods

### *Literature Search*

A primary literature search was conducted on May 26th, 2020 using ISI Web of Science Core Collection with the following search terms: ("effective population size" OR "effective size" OR Ne OR "effective number of breeders" OR Nb) AND (microsatellite* OR allozyme* OR SSR* OR SNP* OR "single nucleotide polymorphism" OR temporal method*). Effective number of breeders ($N_b$) was included in the search because it is an estimate of effective size based on a single breeding season. An additional search was done to include articles that referenced two popular single-sample $N_e$ estimation software packages: LDNe (Waples & Do, 2008), and NeEstimator v2 (Do et al., 2014). Article information was exported into MS Excel, and duplicates were removed, resulting in a total of 4513 articles. The database includes articles published up to the search date of May 26, 2020.

### *Article Screening*

Articles were screened for relevance in two steps, first based on title and abstract, and then based on the full text. If article relevance was unclear at the first step, the article was retained and reviewed further at the second step. For each step, a consistency check was performed to ensure the articles were screened consistently between reviewers (n = 6). All reviewers were given a subset of articles to screen (100 for title and abstract and 20 for full text), and their consistency was determined using a Kappa test (a metric recommended for use in systematic reviews; Collaboration for Environmental Evidence, 2020). The Kappa test gives a measure of how often the reviewers chose the same outcome, with a score of 0 indicating the reviewers agreed no better than expected by chance, and a score of 1 indicating perfect

agreement. In order to proceed with screening, a Kappa score of ≥ 0.6 was necessary to indicate substantial agreement between reviewers.

The articles were screened based on three criteria. (1) Does the study provide an estimate of $N_e$ or $N_b$; (2) Is there a clearly defined, wild animal or plant population; and (3) Does the study use a single-sample genetic estimation method. For criterion (1), only primary literature was included, or where authors obtained DNA sequence information from previous studies and generated a novel estimate of $N_e$ or $N_b$. For criterion (2), populations that were wild, but supplemented from captive breeding or hatchery releases, were not included, except where hatchery introgression was measured to be very low (<5%; an intermediate between two thresholds of 1% and 10% discussed by Allendorf et al., 2004). Populations that were introduced, reintroduced, or translocated were only included if they came from a wild source. Additionally, estimates of $N_e$ for parasites, vectors, and human populations were not included as they did not fit within the scope of the study. To define populations, reviewers primarily used Fst (significance, or values >0.02) and STRUCTURE clusters (Waples & Gaggiotti, 2006). If either of those methods were not available, then reviewers relied on other structuring programs (e.g., BAPS, GENELAND), or relied on assumptions by the authors from previous work in the literature, or assumptions of geographic isolation. Studies where the population definition was vague or not supported by evidence, or where the $N_e$ was estimated above the population level (e.g., if an overall estimate of $N_e$ is provided but Fst and STRUCTURE indicate spatial structure) were not included. Studies where $N_e$ was estimated below the population level (e.g., at sampling sites within a population) were included, but were all combined under the same population ID. For criterion (3), we focused on five main $N_e$ estimation methods, implemented in four software applications (see Table S8). If $N_e$ or $N_b$ was calculated by hand, we ensured that the authors cited

relevant sources for that method (e.g., citing Waples (2006) for the linkage disequilibrium method). We also restricted our estimation methods to only those that used microsatellite, SNP, or allozyme data.

*Data Extraction*

We extracted data from all studies retained after both screening steps and a final consistency check, separate from the initial consistency check during screening. This consistency check involved having all reviewers (n = 6) extract data from the same 20 articles and comparing the information. Any inconsistencies were discussed, and the data extraction protocol was modified to address them. Each line of data entered into the database represents a single estimate from a population. Some populations had multiple estimates over several years, or from different estimation methods, and each of these was entered on a unique row in the database. Data on $N_e$ or $N_c$ were extracted from both tables, and from figures using WebPlotDigitizer software version 4.3 (Rohatgi, 2020). A full list of data extracted is found in Table S9.

*Data Filtering*

After the initial data collation, correction, and organization, there were a total of 8971 $N_e$ estimates (Figure S1). Given that the Linkage Disequilibrium (LD) method comprised 74% of the estimates, we assessed how well the LD method corresponded with estimates from other methods where both were used to estimate $N_e$ or $N_b$ on the same population, to see whether to include multiple estimation methods in our analyses. We found low agreement between the methods using regression analyses ($R^2$ values of <0.1; Figures S3 and S4). Based on this low agreement, and that most data were produced from the LD method, we proceeded with only

using the LD estimates for our analyses. We further filtered the data to remove estimates that could not be used in analyses, e.g., infinite and negative estimates, no sample size reported, or no Waples (2006) bias correction (see figure S2 for more details). We also removed estimates based on replication of markers (i.e., using both SNPs and microsatellites), and software replication (i.e., using both LD$N_e$ and NeEstimator V2). Finally, we created two separate datasets to address spatial and temporal replication (see Table S10 for more information), given that these types of replication could still provide useful information if accounted for with statistical models. The full dataset included spatially and temporally replicated samples, while the filtered dataset removed these two types of replication.

*Mapping and Human Footprint Index*

All populations were mapped in ArcGIS using the coordinates extracted from articles. The maps were created using a World Behrmann equal area projection. For the summary maps, estimates were grouped into 500km×500km grid cells and the count and median of $N_e$ was generated per grid cell. We used the Global Human Footprint dataset (WCS & CIESIN, 2005) to generate a value of human influence for each population at its geographic coordinates. The footprint ranges from zero (no human influence) to 100 (maximum human influence). Values were available in 1km grid cell size and were projected over the point estimates to assign a value of human footprint to each population (i.e., to determine what the footprint value is at the coordinates for each population). The human footprint values were extracted from the map into a spreadsheet to be used for statistical analyses.

*Statistical Analyses*

All statistical analyses were performed in R statistical software version 4.0.5 (R Core Team, 2021). We created generalized linear mixed models (glmms) using the package glmmTMB (Magnusson et al., 2017) to address our three objectives.

To assess taxonomic differences in $N_e$, we created two glmm models, one with $N_e$ and one with $N_b$ as the response variable, both with a gamma distribution and a log link function. Fixed effects for both models included taxonomic group (amphibians, anadromous fishes, birds, freshwater fishes, invertebrates, mammals, marine fishes, and plants), marker type (microsatellite, SNP, or other – allozyme or isozyme), and number of loci (i.e., number of microsatellites or number of SNP, standardized using the z-score transformation). We chose these variables for this analysis as our main interest was to test for differences between taxonomic groups, and we needed to control for differences between marker types and the number of loci used. Fishes were separated into three separate groups (anadromous, freshwater, and marine) due to large sample sizes, and potential differences in life history strategies. Models were also weighted by sample size. Random effects included in the model were the population nested within the study (to account for replication within populations, and any study-level effects). For the models described above, we performed backwards model selection by starting with the fully saturated model described above, removing variables, and comparing the AICc values between models. We chose the model with the lowest AICc, however if models were within 2 AICc of one another, we took the model with the simplest structure (i.e., the most parsimonious). Once the final model was selected, we generated the estimated marginal means of $N_e$ for each taxonomic group using the emmeans package (formerly known as lsmeans; Lenth,

2018), and plotted the means using *ggplot2* (version 3.3.3; Wickham, Chang, & Wickham, 2016).

To assess the agreement with the 50/500 rule, we used $N_e$ values from the dataset without replication to calculate the percentage of populations that fell below the 50 and 500 thresholds. We used the non-replicated dataset to ensure our proportions were not influenced by populations with multiple estimates. We also created models to test whether certain taxonomic groups were more likely to meet 50/500 thresholds than other groups. To do this, we created two new variables, "fifty" and "five hundred" that were coded 0 if the $N_e$ was less than the threshold, and 1 if the $N_e$ met or exceeded the threshold. For these analyses, we only used $N_e$ data, as the thresholds are specific to $N_e$ and not $N_b$, and we again used the non-replicated dataset to be consistent with the proportions generated above. We created two models, one with "fifty" as the response variable, and the other with "five hundred" as the response, both with a binomial distribution. The fixed effects for both models included taxonomic group, marker type, and number of loci. Because we used the non-replicated dataset, the random variable was just study ID. After model selection, we generated the estimated marginal means for the probability of each taxonomic group to meet the threshold.

We created a model to test for a negative relationship between $N_e$ and human footprint and assess how this relationship varied between taxonomic groups. We used human footprint as the response variable, with a gamma distribution with a log link. The predictor variables included taxonomic group, human footprint, the interaction between taxonomic group and human footprint, marker type, and marker number. The random effects were again the population nested within the study. Not all geographic coordinates had a human footprint value associated with them (e.g., in the oceans), so n = 3299 $N_e$ estimates were available in this model (compared to

n = 4145 in the full dataset). Once the final model was selected, we generated values of $N_e$ for each taxonomic group at each value of human footprint between 0 and 100, using the function *emmip* within the *emmeans* package. We then plotted these values in *ggplot2* to visualize the relationship between $N_e$ and human footprint for each taxonomic group.

Finally, to assess differences in $N_e/N_c$ across taxonomic groups, we created a model for the ratio between $N_e$ and $N_c$ or $N_b$ and $N_c$. In this analysis, we combined $N_e$ and $N_b$ data due to a lower sample size of estimates that also included an $N_c$ value. The response variable was the ratio between effective and census size and included a gamma distribution with a log link. Fixed effects were the same as for the $N_e$ and $N_b$ analyses, but also included a variable to indicate whether the ratio was from an $N_e$ or an $N_b$ estimate. Models were weighted by sample size, and random effects were again the population nested within the study.

**Results**

After filtering, the full dataset included 4145 $N_e$ estimates and 827 $N_b$ estimates from 3576 unique populations, extracted from 712 articles. The non-replicated dataset included 3315 $N_e$ estimates and 343 $N_b$ estimates. Our dataset included studies published from 2006 to 2020, with the number of studies published per year increasing steadily (Figure S5). Although the year in which populations were sampled ranged from the mid 1900s to 2019, 92% of estimates came from populations sampled in or after 2000 (Figure S6). Freshwater fishes and reptiles had the most and fewest estimates (n = 1390, n = 211), respectively (see Table S11 for more information on sample sizes). Notably, salmonid fishes comprised 31% of our estimates, and only 6% of the estimates were from populations that had been reintroduced or translocated by humans. Overall, populations sampled were concentrated in Europe and coastal North America, with low

representation in Africa and mid-Western Asia (Figure 2.1). Median $N_e$ values had a hotspot in Alaska, with relatively high values, while low values were spread out across the globe (Figure 2.2).

*Taxonomic Differences in $N_e$ and $N_b$, and Agreement with the 50/500 Rule*

We found strong evidence that both $N_e$ and $N_b$ vary between taxonomic groups. Taxonomic group was significant in all glmms (all $p < 0.001$) and was always included as a predictor variable in the most parsimonious models (Table S12). Marker type was also significant (both $p < 0.001$) and included in the most parsimonious model (Table S12), with microsatellites generating higher estimates than SNPs ($N_e$ = 120.5 and 98.2, respectively, and $N_b$ = 162 and 107, respectively). Comparing $N_e$ across taxonomic groups, marine fishes had the highest $N_e$ at 883.2 and were different from all other groups, with an $N_e$ 3.5x larger than the second-largest group, anadromous fishes ($N_e$ = 255.2). Groups with the lowest $N_e$ were plants, amphibians, and mammals ($N_e$ = 36, 58.5, and 64.1, respectively), while reptiles, invertebrates, freshwater fishes, and birds had intermediate values (Figure 2.3). Comparing $N_b$ values across taxonomic groups, marine fishes again had the largest mean ($N_b$ = 738.6), however they were statistically similar to several other groups (plants, invertebrates, birds, and anadromous fishes; Figure S7). The groups with the lowest $N_b$ were amphibians, reptiles, and mammals ($N_b$ = 48.3, 50.9, and 60.3, respectively).

Many populations do not meet either threshold for the 50/500 rule, and the probability of meeting the thresholds varies between taxonomic groups. Across all taxonomic groups, 38.7% of all populations (1282 of 3315) fell below the threshold of 50, and 85.2% of populations (2825 of 3315) fell below 500. Taxonomic group was significant for both “fifty” and “five hundred”

analyses ($p < 0.001$) and was included as a predictor variable in the most parsimonious model (Table S12). For the five hundred analysis, marker type was also significant ($p = 0.003$); $N_e$ estimates generated using SNPs were more likely to reach the threshold than those estimates generated using microsatellites (probability = 0.17 and 0.07, respectively). Across taxonomic groups, marine fishes are most likely to meet both the 50 and 500 thresholds (probabilities of 0.97 and 0.73, respectively), while all other groups were only likely to meet the 50 threshold or were not likely to meet either threshold (Figure 2.4). Plants, amphibians, and mammals were least likely to meet either threshold, with a ~50% or less chance of reaching an $N_e$ of 50 (probabilities of 0.35, 0.46, and 0.52, respectively).

*Human Footprint*

Our results provide strong support for the hypothesis that human footprint is associated with lower $N_e$, and that the relationship between human footprint and $N_e$ varies between taxonomic groups. The most parsimonious model from this analysis included both human footprint and the interaction between footprint and taxonomic group as fixed effects (both $p < 0.001$; Table S12). Taxonomic group and marker type were also fixed effects in the model ($p < 0.001$), similar to the taxonomic analysis of $N_e$. Across all taxonomic groups, there was a negative relationship between $N_e$ and human footprint ($p < 0.001$, $\chi^2 = 1024.4$), with seven of nine groups showing this negative trend (Figure 2.5). Additionally, the strength of relationship between $N_e$ and human footprint varied between groups. For example, amphibians experienced a 90% decrease in $N_e$ (from 135.9 to 12.6) as human footprint increased from 0 to 100, whereas marine fishes only experienced a 38% decrease (from 775.5 to 481.3). Comparing between taxonomic groups, however, there were two taxonomic groups that had a positive relationship

with human footprint, anadromous fishes and plants (Figure 2.5). Anadromous fishes increased in $N_e$ from 204.2 at a human footprint of 0 to an $N_e$ of 391.1 at a human footprint of 100, and plants increased from $N_e$ of 19.1 to 129.7.

*Ratio Between Effective and Census Size*

Similar to the analysis of $N_e$ and $N_b$, we found strong evidence that the ratio between effective and census size varies between taxa. Taxonomic group was significant ($p < 0.001$) and was included as a predictor variable in the most parsimonious model (Table S12). The type of ratio (i.e., whether it used $N_e$ or $N_b$ alongside $N_c$) was also significant ($p < 0.001$), with $N_b/N_c$ ratios smaller on average than $N_e/N_c$ (0.112 and 0.158, respectively). Anadromous fishes had the highest ratio (0.41), and invertebrates and marine fishes had the lowest ratios (0.009 and 0.03, respectively, Figure 2.6).

**Discussion**

Assessing $N_e$ values across and within taxa on a global scale is important in order to prioritise resource allocation for conservation goals and help forecast adaptive responses of populations to future stressors. Here, we assessed taxonomic differences in $N_e$ and how they related to conservation thresholds and the impact of human activities through a quantitative review. In accordance with our hypotheses, we found consistent taxonomic differences in $N_e$, $N_b$, and the ratio between $N_e$ and $N_c$. Many populations did not meet the 50/500 rule, and the probability of meeting or falling below the thresholds differed based on taxonomic group; amphibians, mammals, and plants were the least likely to meet either threshold. Overall, $N_e$ was negatively correlated with Global Human Footprint, however this relationship differed between

taxonomic groups, and two groups (anadromous fish and invertebrates) were positively correlated with footprint. These results suggest that previous reviews reporting average or median $N_e$ or $N_e/N_c$ were not representative for certain taxonomic groups. Overall, these results may help provide a framework for prioritizing taxa that are least likely to meet thresholds or are most impacted by human activities.

*Taxonomic differences in effective size and the likelihood of meeting the 50/500 rule*

We found strong evidence that there are taxonomic differences in $N_e$ and $N_b$ estimates, with marine fishes having the largest effective sizes, while amphibians, mammals, and plants had the smallest effective sizes. Large $N_e$ in populations of marine species was expected, considering they can inhabit large areas, supporting census sizes of thousands to billions of individuals, with high levels of gene flow between connected populations (Palstra and Ruzzante 2008; Hare et al., 2011; Marandel et al., 2019). Amphibians and mammals, conversely, may be expected to exhibit low $N_e$ values due to their declining populations, and higher risk for extinction compared to other vertebrate species. Ripple et al. (2017) analyzed extinction risk for over 27,000 species of vertebrates and found that large-bodied vertebrates (like marine and terrestrial mammals) were most threatened by direct mortality from humans, while amphibians, the smallest vertebrates, were most threatened by habitat loss and pollution. Amphibians have the highest proportion of listed species out of all vertebrates (41% on the IUCN Red List; Hoffmann et al., 2010), and are specifically at risk from the pathogen *Bd*, which has been found in 42% of amphibian species (Olson et al., 2013). Mammals have 25% of assessed species listed on the IUCN Red List (Hoffmann et al., 2010), and terrestrial mammals (which made up >91% of our mammal $N_e$ estimates) are highly threatened by habitat fragmentation (Crooks et al., 2017). Overall, low $N_e$

values in amphibians and mammals may reflect their declining populations and higher risks for extinction. The taxonomic differences in $N_b$ generally held the same trends as $N_e$, however confidence intervals were much wider, and statistical differences between groups were fewer, likely due to reduced sample sizes (Table S11). Plants had the largest difference between $N_e$ and $N_b$ ($N_e$ = 36, vs $N_b$ = 336), but this may be due to a very low sample size (n = 5). Several explanations could account for the low $N_e$ values in plants, including occurrence of alternate reproductive strategies which may reduce $N_e$ (e.g. selfing or cloning; Orive, 1993; Wright et al., 2013), or declining populations due to climate change, human development, or harvest (39% of assessed vascular plant species are threatened with extinction; Nic Lughadha et al., 2020).

As expected, certain taxonomic groups were more likely to fall below 50/500 thresholds, with the same trends holding true as in the $N_e$ analysis. Of concern, most populations (85.2%) fell below the upper threshold of 500, while over one-third of populations fell below 50 (though these proportions may be slightly inflated, as we removed infinite estimates from our analyses). Palstra and Ruzzante (2008) found that ~70% of published populations fell below 500, while only ~8% fell below 50. Our results show a much larger proportion of populations are at risk of inbreeding in the short-term (i.e., $N_e < 50$) compared to studies included in their 2008 review. In fact, using newer 100/1000 recommendations from Frankham et al. (2014), over half of the populations in this study (56.4%) would fall below $N_e$ of 100, and over 90% would fall below 1000. While we expect that these proportions may be slightly over-estimated (due to removal of infinite estimates and potential for downward biases in $N_e$ estimates, discussed below), it is still very likely that many wild populations do not meet the thresholds for either the 50/500 or 100/1000 rules. Although life history traits likely play a role in determining whether a population meets these thresholds (e.g., in marine fishes), some taxonomic groups seem to be more at risk

than others of low $N_e$ and therefore extinction. Our results here can help identify groups most at-risk of falling below critical thresholds (e.g., amphibians, mammals, and plants) to better prioritize conservation goals and manage declining populations.

*The influence of human footprint on $N_e$ varies across taxonomic groups*

According to our predictions we found that across all populations $N_e$ was negatively correlated with human footprint, however this relationship varied between taxonomic groups in both strength and direction. The human footprint gives a measure of human land use, population density, and infrastructure (e.g., built environments and roads; WCS & CIESIN, 2005), all of which can contribute to habitat degradation or fragmentation, and may result in reduced population sizes or $N_e$. Several studies have shown decreases in $N_e$ associated with anthropogenic habitat fragmentation in several different taxa (Alò & Turner, 2005; Browne & Karubian, 2018; Keller et al., 2005; Sumner et al., 2004). The overall relationship with $N_e$ and human footprint, while negative, was relatively weak (with a coefficient of -0.024), which is likely due to the difficulty in generalizing across a wide variety of species (some of which may benefit from human influences).

Within taxonomic groups, amphibians and mammals had the largest decreases in $N_e$ as footprint increased from 0 to 100 (90.8% and 84.3% decreases, respectively). Contrary to our results, Schmidt and Garroway (2021) found a weak but positive relationship between $N_e$ and human footprint in amphibians in North America, and generally found minimal evidence of an effect of urbanization on amphibian genetics. On the other hand, the impacts of human footprint on mammal populations seems consistent in other studies, with high human footprint values being associated with extinction, anthropogenic mortality, and a negative relationship with $N_e$

(Di Marco et al., 2018; J. E. Hill et al., 2020; Schmidt et al., 2021). Interestingly, for plants and anadromous fishes, $N_e$ was positively correlated with footprint. In plants, this relationship may be due to positive influences of human activities, like transportation networks or agricultural development, on dispersal and gene flow (Auffret et al., 2014; Bullock & Pufal, 2020). For anadromous fishes, there may be a positive influence of humans from habitat restoration or facilitation of migration (e.g. salmon ladders), but more likely this relationship is due to limitations in capturing the true location of anadromous populations due to their migratory nature. The human footprint index only provides values for terrestrial biomes, with no data on marine environments (hence the low sample size for marine fishes); therefore, most footprint values associated with anadromous fishes are from populations sampled in their freshwater environment, which may not represent the conditions experienced throughout their life cycle. Many human settlements also occur along large bodies of water which are likely able to support large populations of anadromous fishes; therefore, there may be a covariance between human settlements and rivers that have populations with large $N_e$. While this caveat is especially pertinent for anadromous fish, we note that other populations that may inhabit large ranges (e.g., large mammals or other migratory species) could also be misrepresented in the point estimate of human footprint based on where they were sampled.

*Effective to census size ratios reveal taxonomic trends and uncertainties*

Effective to census size ratios varied widely between taxonomic groups; marine fishes and invertebrates had very small ratios (<0.05) and anadromous fishes had the largest ratio. In general, taxonomic differences in $N_e/N_c$ are expected, due to the fact that life history traits play an important role in determining the relationship between $N_e$ and N (Waples et al., 2013). The

occurrence of very low $N_e/N$ ratios has been discussed in the literature since the early 1990s, and it is generally accepted that they occur due to very high fecundities and high juvenile mortalities, such as in species with Type III survivorship curves like marine organisms (Hedgecock, 1994; Hedrick, 2005; Waples, 2002b). Indeed, several papers have reported very low $N_e/N$ for marine fish populations (e.g., Hauser et al., 2002; Hoarau et al., 2005; Turner et al., 2002). Interestingly, invertebrates had a lower ratio than even marine fishes, which could be explained by the fact that 50% of the ratio estimates from invertebrates come from marine molluscs, which are known to be broadcast spawners with high fecundity and high juvenile mortality (Ramirez Llodra, 2002). However, there are two important caveats to these extremely low $N_e/N$ ratios in marine fishes and invertebrates. Firstly, these two groups have the smallest sample sizes (n = 9 and n = 6, respectively; Table S11), so it is hard to make conclusions based on these data. Additionally, it is important to note that very small $N_e/N_c$ ratios in marine species could be due to an under-estimation of the true $N_e$; Waples (2016) showed that in marine fish populations with a high $N_e$ (e.g., $10^6$), the linkage disequilibrium method can underestimate $N_e$ by up to three or four orders of magnitude.

Anadromous fishes had a high $N_e/N$, which could be in part due to the high proportion of Atlantic salmon estimates making up this category (62%). In Atlantic salmon, mature male parr can take part in mating, and therefore contribute to the $N_e$, but are often not counted as part of the $N_c$. This can result in an over-estimation of $N_e/N_c$ (see Johnstone et al., 2013; Perrier et al., 2014; Saura et al., 2008), which may be inflating the value in our analysis. We found, similar to Palstra and Fraser (2012), that many studies did not report uncertainty around their $N_c$ estimates, which may reflect decreased accuracy in these values. Had we chosen to remove these estimates, however, it would have considerably reduced our sample size (from n = 537 to n = 167). Instead,

we chose to retain all $N_c$ estimates to maximise our sample size, with the caveat that some $N_c$ estimates may be less accurate than others. It remains extremely important for authors to report uncertainty around both $N_e$ and $N_c$, but especially $N_c$. We did not quantify here the number of incorrectly linked $N_e$ and $N_c$ ratios, but it is likely that there are still many studies being published without proper knowledge of linking $N_e$ and $N_c$ (see Waples, 2005, and Palstra and Fraser 2012).

*Potential downward bias in $N_e$ estimates*

In this study, we performed a much-needed review of the effective size literature, with strict filtering criteria, extensive georeferencing, and representation from across the globe. However, there is still an important caveat to acknowledge when drawing conclusions from these results. Namely, that the $N_e$ estimates here are likely biased downward as a result of several factors. First, we removed any estimates that were infinite or negative, as these could not be incorporated into our analytical results. Some of these populations could have been large, and the sample size used was insufficient to generate a point-estimate of $N_e$ (Do et al., 2014). Therefore, by removing the negative or infinite estimates, we likely removed data on large populations that could have increased our averages by taxonomic group. We do, however, believe that the trends between groups would remain the same, as most taxonomic groups had the same proportion of infinite values (~10-15%; marine fishes and invertebrates had a higher proportion, ~30%, but these two groups were already on the "larger" end of the $N_e$ spectrum between groups). Secondly, the LD method is known to have a hard time reliably estimating large $N_e$ values, unless the sample sizes are sufficiently large (Waples & Do, 2010). We did control for this in our analyses by using sample size as weight, with larger sample sizes being weighted more than

those with small sample sizes. Nevertheless, there was still likely downward bias in our analyses. For large marine populations, around 1% of the total individuals may need to be sampled to ensure accurate estimates (Marandel et al., 2019), which could require sample sizes in the thousands or millions (the largest sample size in our database was 5413). The final source of potential downward bias in our dataset, is the sampling bias associated with which populations are studied or reported in the literature. At-risk populations with small or declining $N_c$ and $N_e$ may be more likely to be studied due to funding from governmental agencies (e.g. the U.S. Fish and Wildlife Service, which manages the Endangered Species Act; Mahoney, 2009), or due to more interest in published results for threatened species. Overall, although the $N_e$ estimates reported here are likely biased downward, we believe the taxonomic differences still hold true, and these results are still representative of which taxonomic groups are most at risk of falling below conservation thresholds, and which are most impacted by human activities.

*Conclusions and recommendations*

Our study revealed important taxonomic trends in effective size estimates and their relationship with the 50/500 rule and the human footprint. Many populations did not reach critical conservation thresholds, and generally $N_e$ was negatively associated with human footprint. We identified certain taxonomic groups (i.e., amphibians and mammals) most at risk of falling below these thresholds, potentially due to the negative impacts of human activities. These results can help guide future recommendations for species listing criteria or prioritize assessment of certain taxa most at risk of falling below conservation thresholds. However, these results have also generated several important questions to be addressed in future works:

(1) Do these relationships hold true for other single-sample estimators of $N_e$, such as the sibship frequency or Bayesian computation methods? Our analyses showed low agreement between estimators, likely a result of the different methods estimating different types of $N_e$ (the sibship method estimates inbreeding $N_e$, while LD and Bayesian likely estimate a combination of inbreeding and variance $N_e$; Wang, 2016). For now, we should take caution in extending the conclusions from this review to other estimators of contemporary $N_e$ until we know more about how they relate to one another.

(2) Are there finer-scale differences in $N_e$ between taxonomic groups (e.g., family, genus, or species)? We limited our scope to broad taxonomic groupings of populations, but the data exist from certain genera or species (e.g., brook trout or Atlantic salmon) to generate much finer-scale estimates of $N_e$ and further explore the influence of human footprint on effective size. Studies on the relationship between genetic diversity and urbanization or anthropogenic fragmentation have shown that individual species or genera can vary in the strength and direction of their responses (Habrich et al., 2021; Schmidt et al., 2021). We already showed taxonomic variation in response to human footprint, but finer-scale analyses are likely to be more accurate.

(3) How should we apply the 50/500 or 100/1000 rule to varying taxa? The 50/500 and 100/1000 rules are generalizations based on shared responses across taxa and predictions based on inbreeding depression, fitness, and population size (Frankham et al., 2014). The results from our study reveal that some taxa may exist at much larger $N_e$ than others (i.e., marine fishes), and therefore the 50 and 500 thresholds may be less appropriate for these taxa. Future work should focus on disentangling whether the differences seen here between taxonomic groups is due to declining populations (e.g., in amphibians and mammals), or due to life history characteristics resulting in differing $N_e$ values.

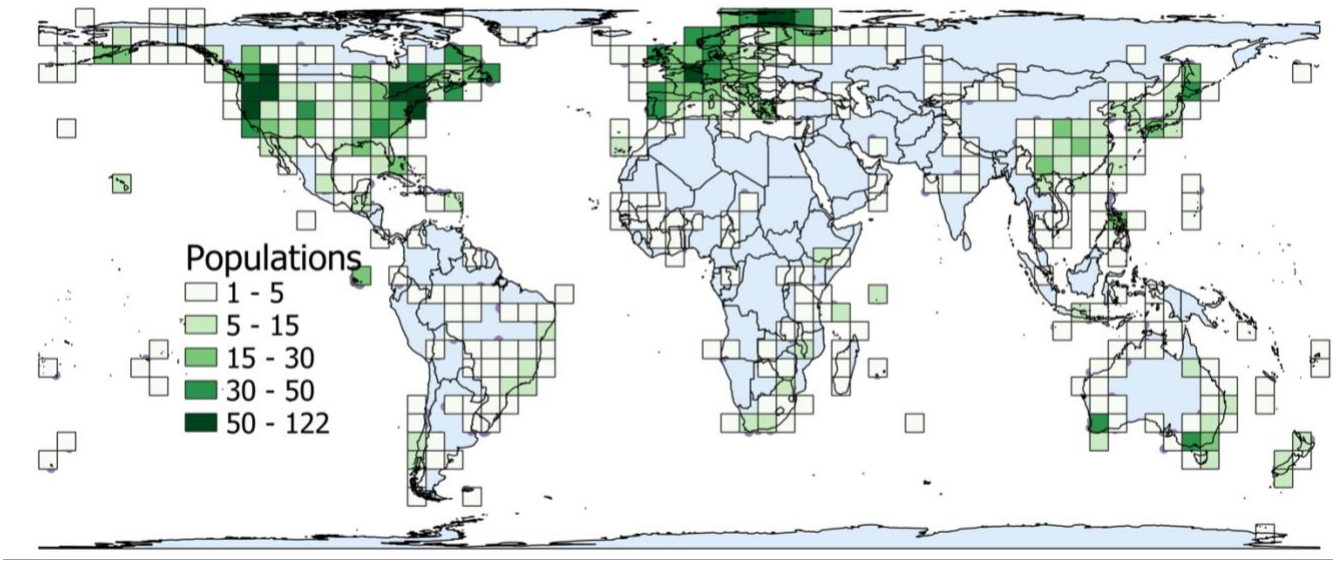


**Figure 2.1**: Global map showing the number of populations sampled in each 500×500km grid cell. Data is projected with the world Behrmann projection

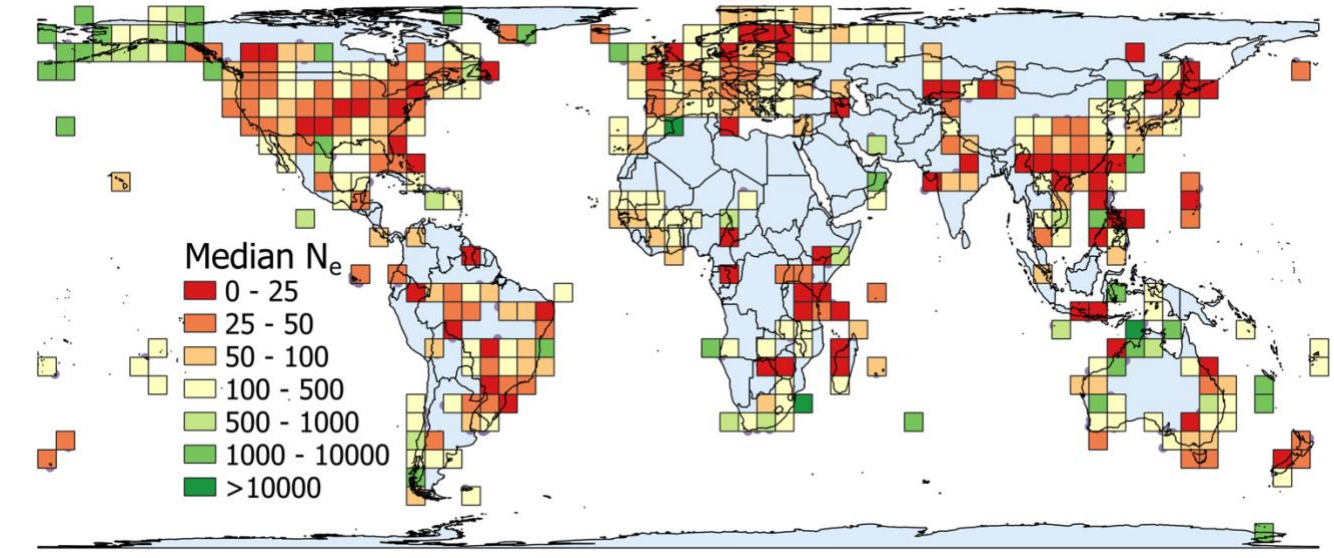


**Figure 2.2**: Global map showing the median $N_e$ value across all taxonomic groups in each 500×500km grid cell. Data is projected with the world Behrmann projection

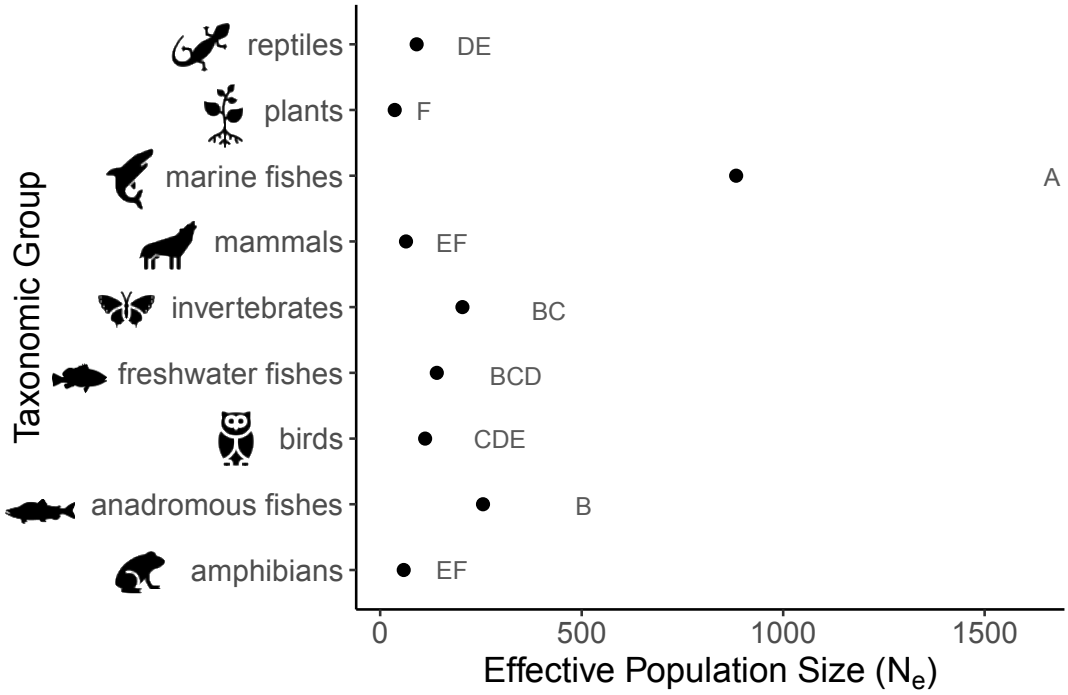


**Figure 2.3**: Mean effective population size ($N_e$) across taxonomic groups, accounting for differences in marker type. Shaded green bars represent the 95% confidence intervals. Groups with no shared letters are statistically different from one another.

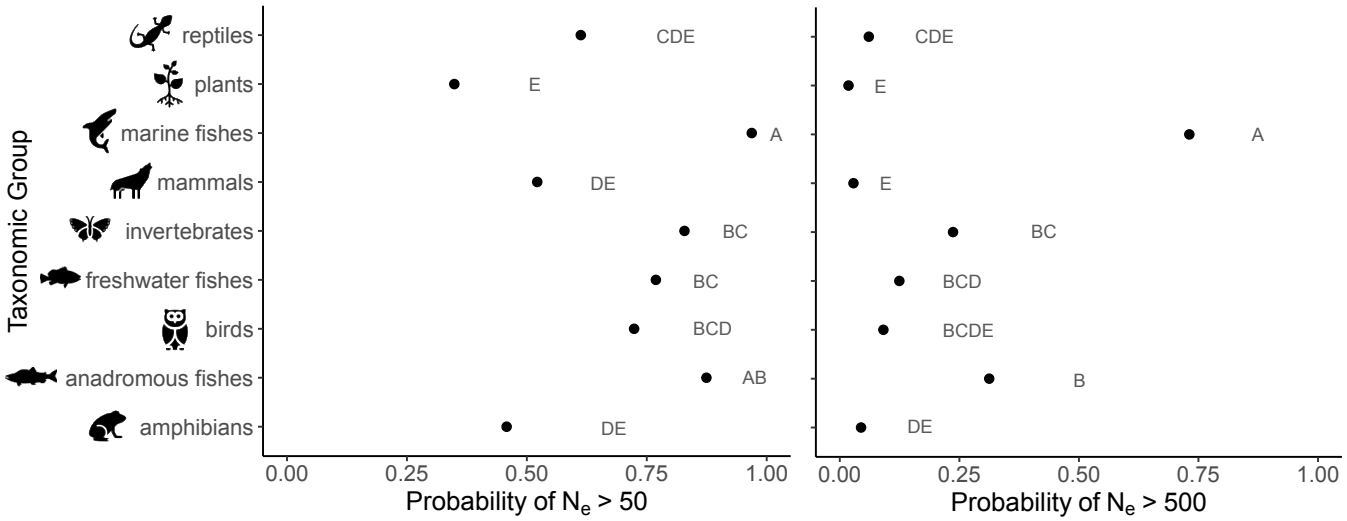


**Figure 2.4**: Probabilities of reaching an effective population size ($N_e$) of a) 50 or higher, and b) 500 or higher across taxonomic groups. Shaded green bars represent the 95% confidence intervals. Groups with no shared letters are statistically different from one another. Models were generated using the non-replicated dataset.

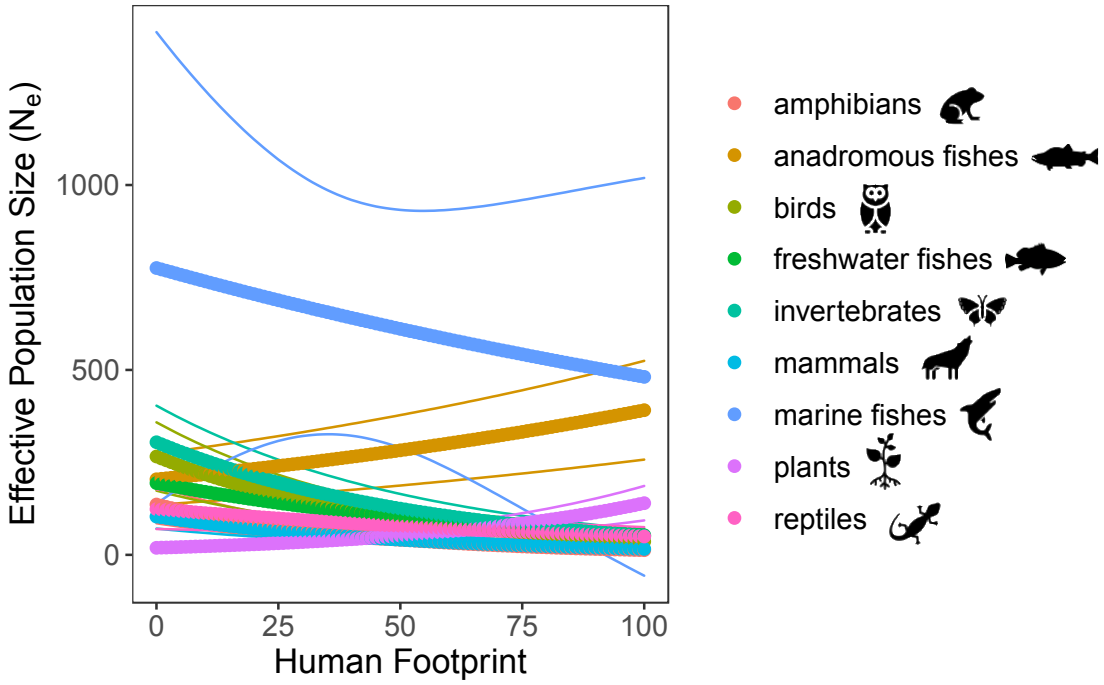


**Figure 2.5**: Relationship between effective population size ($N_e$) and human footprint across taxonomic groups. Shaded areas represent 95% confidence intervals.

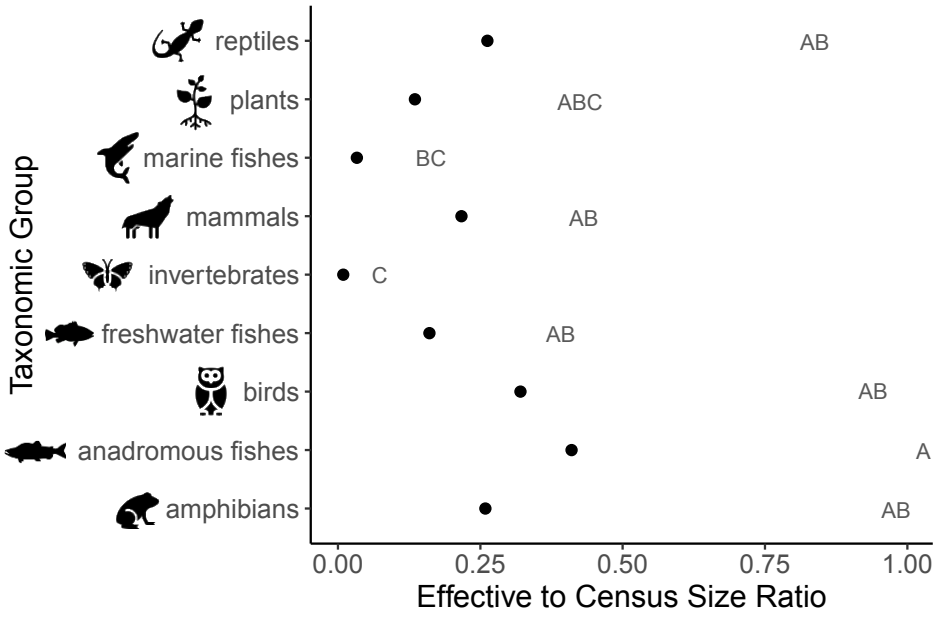


**Figure 2.6**: Effective to census size ratio ($N_e/N_c$ and $N_b/N_c$) across taxonomic groups, accounting for differences between $N_e$ and $N_b$. Shaded green bars represent the 95% confidence intervals. Groups with no shared letters are statistically different from one another.

**References**


Allendorf, F.W., Leary, R.F., Hitt, N.P., Knudsen, K.L., Lundquist, L.L., & Spruell, P. Intercrosses and the U.S. Endangered Species Act: Should Hybridized Populations be Included as Westslope Cutthroat Trout? *Conservation Biology, 18*(5), 1203–1213. https://doi.org/10.1111/j.1523-1739.2004.00305.x

Alò, D., & Turner, T. F. (2005). Effects of Habitat Fragmentation on Effective Population Size in the Endangered Rio Grande Silvery Minnow. *Conservation Biology*, *19*(4), 1138–1148. https://doi.org/10.1111/j.1523-1739.2005.00081.x

Auffret, A. G., Berg, J., & Cousins, S. A. O. (2014). The geography of human-mediated dispersal. *Diversity and Distributions*, *20*(12), 1450–1456. https://doi.org/10.1111/ddi.12251

Barbato, M., Orozco-terWengel, P., Tapio, M., & Bruford, M. W. (2015). SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, *6*, 109.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity: Biodiversity and climate change. *Ecology Letters*, *15*(4), 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Bijlsma, R., Bundgaard, J., & Boerema, A. C. (2000). Does inbreeding affect the extinction risk of small populations?: Predictions from Drosophila. *Journal of Evolutionary Biology*, *13*(3), 502–514.

Browne, L., & Karubian, J. (2018). Habitat loss and fragmentation reduce effective gene flow by disrupting seed dispersal in a neotropical palm. *Molecular Ecology*, *27*(15), 3055–3069. https://doi.org/10.1111/mec.14765

Bullock, J. M., & Pufal, G. (2020). Human-mediated dispersal as a driver of vegetation dynamics: A conceptual synthesis. *Journal of Vegetation Science*, *31*(6), 943–953. https://doi.org/10.1111/jvs.12888

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*(5), e1400253.

Collaboration for Environmental Evidence. (2020). *Guidelines for systematic review and evidence synthesis in environmental Management: Eligibility Screening*. http://www.environmentalevidence.org/guidelines/section-6

Crooks, K. R., Burdett, C. L., Theobald, D. M., King, S. R. B., Di Marco, M., Rondinini, C., & Boitani, L. (2017). Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. *Proceedings of the National Academy of Sciences*, *114*(29), 7635–7640. https://doi.org/10.1073/pnas.1705769114

Crow, J. F., & Kimura, M. (1970). *An introduction to population genetics theory*. Harper & Row, New York.

Di Marco, M., Venter, O., Possingham, H. P., & Watson, J. E. M. (2018). Changes in human footprint drive changes in species extinction risk. *Nature Communications*, *9*(1), 4621. https://doi.org/10.1038/s41467-018-07049-5

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 (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Dunham, J., Peacock, M., Tracy, C. R., Nielsen, J., & Vinyard, G. (1999). Assessing extinction risk: Integrating genetic information. *Conservation Ecology*, *3*(1).

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. *Genetical Research*, *66*(2), 95–107. https://doi.org/10.1017/S0016672300034455

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frankham, R., Bradshaw, C. J. A., & Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, *170*, 56–63. https://doi.org/10.1016/j.biocon.2013.12.036

Franklin, I. R. (1980). Evolutionary change in small populations. *Conservation Biology: An Evolutionary-Ecological Perspective*, *395*.

García-Dorado, A. (2015). On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. *Heredity*, *115*(3), 185–187. https://doi.org/10.1038/hdy.2015.28

Habrich, A. K., Lawrence, E. R., & Fraser, D. J. (2021). Varying genetic imprints of road networks and human density in North American mammal populations. *Evolutionary Applications*, *14*(6), 1659–1672. https://doi.org/10.1111/eva.13232

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg, K., & Palstra, F. (2011). Understanding and Estimating Effective Population Size for Practical Application in Marine Species Management: Applying Effective Population Size Estimates to Marine Species Management. *Conservation Biology*, *25*(3), 438–449. https://doi.org/10.1111/j.1523-1739.2010.01637.x

Hauser, L., Adcock, G. J., Smith, P. J., Bernal Ramirez, J. H., & Carvalho, G. R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (Pagrus auratus). *Proceedings of the National Academy of Sciences*, *99*(18), 11742–11747. https://doi.org/10.1073/pnas.172242899

Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms? *Genetics and Evolution of Aquatic Organisms*, 122–134.

Hedrick, P. (2005). LARGE VARIANCE IN REPRODUCTIVE SUCCESS AND THE Ne/N RATIO. *Evolution*, *59*(7), 1596–1599. https://doi.org/10.1111/j.0014-3820.2005.tb01809.x

Hill, J. E., DeVault, T. L., Wang, G., & Belant, J. L. (2020). Anthropogenic mortality in mammals increases with the human footprint. *Frontiers in Ecology and the Environment*, *18*(1), 13–18. https://doi.org/10.1002/fee.2127

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, *38*(3), 209–216. https://doi.org/10.1017/S0016672300020553

Hoarau, G., Boon, E., Jongma, D. N., Ferber, S., Palsson, J., Van der Veer, H. W., Rijnsdorp, A. D., Stam, W. T., & Olsen, J. L. (2005). Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice ( *Pleuronectes platessa* L.). *Proceedings of the Royal Society B: Biological Sciences*, *272*(1562), 497–503. https://doi.org/10.1098/rspb.2004.2963

Hoban, S., Bruford, M., D'Urban Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P. A., Paz-Vinas, I., Sjögren-Gulve, P., Segelbacher, G., Vernesi, C., Aitken, S., Bertola, L. D., Bloomer, P., Breed, M., Rodríguez-Correa, H., Funk, W. C., Grueber, C. E., Hunter, M. E., Jaffe, R., … Laikre, L. (2020). Genetic diversity targets and indicators in the CBD

post-2020 Global Biodiversity Framework must be improved. *Biological Conservation*, *248*, 108654. https://doi.org/10.1016/j.biocon.2020.108654

Hoffmann, M., Hilton-Taylor, C., Angulo, A., Böhm, M., Brooks, T. M., Butchart, S. H. M., Carpenter, K. E., Chanson, J., Collen, B., Cox, N. A., Darwall, W. R. T., Dulvy, N. K., Harrison, L. R., Katariya, V., Pollock, C. M., Quader, S., Richman, N. I., Rodrigues, A. S. L., Tognelli, M. F., … Stuart, S. N. (2010). The Impact of Conservation on the Status of the World's Vertebrates. *Science*, *330*(6010), 1503–1509. https://doi.org/10.1126/science.1194442

Jamieson, I. G., & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution*, *27*(10), 578–584. https://doi.org/10.1016/j.tree.2012.07.001

Johnstone, D. L., O'Connell, M. F., Palstra, F. P., & Ruzzante, D. E. (2013). Mature male parr contribution to the effective size of an anadromous Atlantic salmon ( *Salmo salar* ) population over 30 years. *Molecular Ecology*, *22*(9), 2394–2407. https://doi.org/10.1111/mec.12186

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

Keller, I., Excoffier, L., & Largiader, C. R. (2005). Estimation of effective population size and detection of a recent population decline coinciding with habitat fragmentation in a ground beetle. *Journal of Evolutionary Biology*, *18*(1), 90–100. https://doi.org/10.1111/j.1420-9101.2004.00794.x

Lenth, R. (2018). Package 'lsmeans.' *The American Statistician*, *34*(4), 216–221.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J., & Stuart, S. N. (2008). Quantification of Extinction Risk: IUCN's System for Classifying Threatened Species. *Conservation Biology*, *22*(6), 1424–1442. https://doi.org/10.1111/j.1523-1739.2008.01044.x

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., van Bentham, K., Bolker, B., Brooks, M., & Brooks, M. M. (2017). Package 'glmmTMB.' *R Package Version 0.2. 0*.

Mahoney, J. (2009). *What Determines the Level of Funding for an Endangered Species? 11*, 18.

Marandel, F., Lorance, P., Berthelé, O., Trenkel, V. M., Waples, R. S., & Lamy, J.-B. (2019). Estimating effective population size of large marine populations, is it feasible? *Fish and Fisheries*, *20*(1), 189–198. https://doi.org/10.1111/faf.12338

Nic Lughadha, E., Bachman, S. P., Leão, T. C. C., Forest, F., Halley, J. M., Moat, J., Acedo, C., Bacon, K. L., Brewer, R. F. A., Gâteblé, G., Gonçalves, S. C., Govaerts, R., Hollingsworth, P. M., Krisai-Greilhuber, I., Lirio, E. J., Moore, P. G. P., Negrão, R., Onana, J. M., Rajaovelona, L. R., … Walker, B. E. (2020). Extinction risk and threats to plants and fungi. *PLANTS, PEOPLE, PLANET*, *2*(5), 389–408. https://doi.org/10.1002/ppp3.10146

Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample: Estimation of effective number of breeders. *Evolutionary Applications*, *1*(3), 462–474. https://doi.org/10.1111/j.1752-4571.2008.00015.x

O'Dea, R. E., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W. A., Parker, T. H., Gurevitch, J., Page, M. J., Stewart, G., Moher, D., & Nakagawa, S. (2021). Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: A PRISMA extension. *Biological Reviews*, *96*(5), 1695–1722. https://doi.org/10.1111/brv.12721

Olson, D. H., Aanensen, D. M., Ronnenberg, K. L., Powell, C. I., Walker, S. F., Bielby, J., Garner, T. W. J., Weaver, G., The Bd Mapping Group, & Fisher, M. C. (2013). Mapping the Global Emergence of Batrachochytrium dendrobatidis, the Amphibian Chytrid Fungus. *PLoS ONE*, *8*(2), e56802. https://doi.org/10.1371/journal.pone.0056802

Orive, M. E. (1993). Effective population size in organisms with complex life-histories. *Theoretical Population Biology*, *44*(3), 316–340.

Palstra, F. P., & Fraser, D. J. (2012). Effective/census population size ratio estimation: A compendium and appraisal. *Ecology and Evolution*, *2*(9), 2357–2365. https://doi.org/10.1002/ece3.329

Palstra, F. P., & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, *17*(15), 3428–3447. https://doi.org/10.1111/j.1365-294X.2008.03842.x

Perrier, C., Normandeau, É., Dionne, M., Richard, A., & Bernatchez, L. (2014). Alternative reproductive tactics increase effective population size and decrease inbreeding in wild Atlantic salmon. *Evolutionary Applications*, *7*(9), 1094–1106. https://doi.org/10.1111/eva.12172

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187).

Pudovkin, A. I., Zaykin, D. V., & Hedgecock, D. (1996). On the Potential for Estimating the Effective Number of Breeders From Heterozygote-Excess in Progeny. *Genetics*, *144*(1), 383–387. https://doi.org/10.1093/genetics/144.1.383

R Core Team. (2021). *R: A language and environment for statistical computing.* (4.0.5) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramirez Llodra, E. (2002). Fecundity and life-history strategies in marine invertebrates. *Advances in Marine Biology*, *43*, 87–170. https://doi.org/10.1016/S0065-2881(02)43004-0

Ripple, W. J., Wolf, C., Newsome, T. M., Hoffmann, M., Wirsing, A. J., & McCauley, D. J. (2017). Extinction risk is most acute for the world's largest and smallest vertebrates. *Proceedings of the National Academy of Sciences*, *114*(40), 10678–10683. https://doi.org/10.1073/pnas.1702078114

Rohatgi, A. (2020). *WebPlotDigitizer Version 4.3*. https://automeris.io/WebPlotDigitizer

Rosser, A. M., & Mainka, S. A. (2002). Overexploitation and Species Extinctions. *Conservation Biology*, *16*(3), 584–586. https://doi.org/10.1046/j.1523-1739.2002.01635.x

Saura, M., Caballero, A., Caballero, P., & Morán, P. (2008). Impact of precocious male parr on the effective size of a wild population of Atlantic salmon. *Freshwater Biology*, *53*(12), 2375–2384. https://doi.org/10.1111/j.1365-2427.2008.02062.x

Schmidt, C., Domaratzki, M., Kinnunen, R. P., Bowman, J., & Garroway, C. J. (2021). Continent-wide effects of urbanization on bird and mammal genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*, 8.

Schmidt, C., & Garroway, C. J. (2021). The population genetics of urban and rural amphibians in North America. *Molecular Ecology*, *30*(16), 3918–3929. https://doi.org/10.1111/mec.16005

Sumner, J., Jessop, T., Paetkau, D., & Moritz, C. (2004). Limited effect of anthropogenic habitat fragmentation on molecular diversity in a rain forest skink, Gnypetoscincus queenslandiae: FRAGMENTATION AND MOLECULAR DIVERSITY IN A SKINK. *Molecular Ecology*, *13*(2), 259–269. https://doi.org/10.1046/j.1365-294X.2003.02056.x

Tallmon, D. A., Koyuk, A., Luikart, G., & Beaumont, M. A. (2008). COMPUTER PROGRAMS: Onesamp: a program to estimate effective population size using approximate Bayesian computation: COMPUTER PROGRAMS. *Molecular Ecology Resources*, *8*(2), 299–301. https://doi.org/10.1111/j.1471-8286.2007.01997.x

Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, *118*(10), e2015096118. https://doi.org/10.1073/pnas.2015096118

Turner, T. F., Wares, J. P., & Gold, J. R. (2002). Genetic Effective Size Is Three Orders of Magnitude Smaller Than Adult Census Size in an Abundant, Estuarine-Dependent Marine Fish ( *Sciaenops ocellatus* ). *Genetics*, *162*(3), 1329–1339. https://doi.org/10.1093/genetics/162.3.1329

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., & Watson, J. E. M. (2016). Global terrestrial Human Footprint maps for 1993 and 2009. *Scientific Data*, *3*(1), 160067. https://doi.org/10.1038/sdata.2016.67

Wang, J. (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, *18*(10), 2148–2164. https://doi.org/10.1111/j.1365-294X.2009.04175.x

Wang, J. (2016). A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*, *25*(19), 4692–4711.

Waples, R. S. (2002a). Definition and estimation of effective population size in the conservation of endangered species. *Population Viability Analysis*, 147–168.

Waples, R. S. (2002b). Evaluating the effect of stage-specific survivorship on the $N_e / N$ ratio. *Molecular Ecology*, *11*(6), 1029–1037. https://doi.org/10.1046/j.1365-294X.2002.01504.x

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, *7*(2), 167–184. https://doi.org/10.1007/s10592-005-9100-y

Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–756. https://doi.org/10.1111/j.1755-0998.2007.02061.x

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

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419–1439. https://doi.org/10.1111/j.1365-294X.2006.02890.x

Waples, R. S., Luikart, G., Faulkner, J. R., & Tallmon, D. A. (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1768), 20131339. https://doi.org/10.1098/rspb.2013.1339

WCS, W. C. S., & CIESIN, C. for I. E. S. I. N.-C. U. (2005). *Last of the Wild Project, Version 2, (LWP-2): Global Human Footprint Dataset (Geographic)*. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4M61H5F

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2.' *Create Elegant Data Visualisations Using the Grammar of Graphics. Version*, *2*(1), 1–189.

Wilson, M. C., Chen, X.-Y., Corlett, R. T., Didham, R. K., Ding, P., Holt, R. D., Holyoak, M., Hu, G., Hughes, A. C., Jiang, L., Laurance, W. F., Liu, J., Pimm, S. L., Robinson, S. K., Russo, S. E., Si, X., Wilcove, D. S., Wu, J., & Yu, M. (2016). Habitat fragmentation and biodiversity conservation: Key findings and future challenges. *Landscape Ecology*, *31*(2), 219–227. https://doi.org/10.1007/s10980-015-0312-3

Wright, S. I., Kalisz, S., & Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1760), 20130133. https://doi.org/10.1098/rspb.2013.0133

Yates, M. C., Bowles, E., & Fraser, D. J. (2019). Small population size and low genomic diversity have no effect on fitness in experimental translocations of a wild fish. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1916), 20191989. https://doi.org/10.1098/rspb.2019.1989

Zhdanova, O. L., & Pudovkin, A. I. (2008). Nb_HetEx: A Program to Estimate the Effective Number of Breeders. *Journal of Heredity*, *99*(6), 694–695. https://doi.org/10.1093/jhered/esn061

**General Conclusions**

The effective population size ($N_e$) is a key genetic estimate that gives a measure of the viability of a population in both the short- and long-term (Franklin, 1980; Waples, 2002). Although $N_e$ is widely reported in the literature, there still lacks effective integration of $N_e$ into monitoring programs, and there is still much we do not understand about how it changes temporally within populations exposed to various stressors. Throughout the two chapters of this thesis, I attempted to bridge these gaps by providing empirical evidence on how effective size changes within populations exposed to size-selective harvest (Chapter 1), and by summarizing $N_e$ estimates in wild populations and identifying areas for future focus within research and policy (Chapter 2).

In Chapter 1, I showed how, in populations of brook trout exposed to intense size-selective harvest, $N_b$ can be resilient over the short-term due to density-dependent compensation, even with substantial decreases in $N_c$. Understanding how effective size responds to stressors like exploitation can help guide genetic monitoring in fisheries and other disciplines. Although the scope of Chapter 1 was relatively narrow, it demonstrated how important experimental studies are in parsing out trends in wild populations that can undergo natural fluctuations in census and effective size. My second chapter focused more widely on a comprehensive review and analysis of effective size in wild populations, and found that taxonomic groups differed in $N_e$, $N_b$, and $N_e/N_c$, and their responses to human impacts. This work showed how generalizing across taxa can result in over- or under-estimation of conservation targets and highlighted how human influences on wild populations can be very nuanced. The results from this work can act as a reference for future empirical research, to generate realistic expectations of what $N_e$ or $N_e/N_c$ can look like in populations, and it can help inform prioritization of the most impacted taxa.

Broadly speaking, one of the important conclusions from these works was that many wild populations do not meet conservation thresholds set out in the literature (i.e. the 50/500 rule). In Chapter 2, I showed how over one-third of reported $N_e$ estimates in the literature fell below an $N_e$ of 50, and over 85% fell below 500. Extending these criteria to the $N_{e(adj)}$ calculated from Chapter 1, all of the populations fell below 500 for all time periods, and four fell below 50 at least once during the three years of monitoring (all three harvest lakes, plus McNair). As I suggest earlier, there may be caveats in using generalized thresholds like the 50/500 rule, that ignore important life history adaptations of specific species that may allow them to exist at generally higher or lower $N_e$ in wild populations. One of the key areas for future research is in exploring genera- or species-specific trends in $N_e$ and $N_e/N_c$ and implementing them in monitoring or risk assessment decisions.

Another important topic brought up throughout my thesis was the interaction between genetic and demographic factors in wild populations. While $N_e$ is a genetic measure, it is heavily influenced by demographic variables such as census size, breeding systems, or the sex ratio (Franklin, 1980). In Chapter 1, this was shown through the genetic compensation experienced in the harvest lakes; change in $N_b$ was buffered through density-dependent processes during mating (i.e., demographic factors had an impact on $N_b$). This relationship was also touched on in Chapter 2, when discussing the influence of life history traits on $N_e$ and its relationship with $N_c$; in order to predict how $N_e$ may change in threatened populations, it is imperative to understand the demographics and life history of that population. My thesis highlights the importance of integrating both genetic and demographic factors into research, monitoring, and policy.

**References**

Alò, D., and T. F. Turner. 2005. Effects of Habitat Fragmentation on Effective Population Size in the Endangered Rio Grande Silvery Minnow. Conservation Biology 19(4):1138–1148.

Bellard, C., C. Bertelsmeier, P. Leadley, W. Thuiller, and F. Courchamp. 2012. Impacts of climate change on the future of biodiversity: Biodiversity and climate change. Ecology Letters 15(4):365–377.

Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in conservation biology. Springer Science & Business Media.

Burgman, M. A., and D. Yemshanov. 2013. Risks, decisions and biological conservation. Diversity and Distributions 19(5–6):485–489.

Chavez-Pesqueira, M., P. Suarez-Montes, G. Castillo, and J. Nunez-Farfan. 2014. Habitat fragmentation threatens wild populations of Carica papaya (Caricaceae) in a lowland rainforest. American Journal of Botany 101(7):1092–1101.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.

Durrant, C. J., J. R. Stevens, C. Hogstrand, and N. R. Bury. 2011. The effect of metal pollution on the population genetic structure of brown trout (Salmo trutta L.) residing in the River Hayle, Cornwall, UK. Environmental Pollution 159(12):3595–3603.

Frankham, R. 2003. Genetics and conservation biology. Comptes Rendus Biologies 326:22–29.

Frankham, R. 2005. Genetics and extinction. Biological Conservation 126(2):131–140.

Frankham, R., C. J. A. Bradshaw, and B. W. Brook. 2014. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biological Conservation 170:56–63.

Franklin, I. R. 1980. Evolutionary change in small populations. Pages 135–149 *in* M. E. Soule and B. A. Wilcox, editors. Conservation Biology - An Evolutionary-Ecological Perspective.

García-Dorado, A. 2015. On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. Heredity 115(3):185–187.

Harwood, J. 2000. Risk assessment and decision analysis in conservation. Biological Conservation 95(2):219–226.

Jamieson, I. G., and F. W. Allendorf. 2012. How does the 50/500 rule apply to MVPs? Trends in Ecology & Evolution 27(10):578–584.

Mirimin, L., B. Macey, S. Kerwath, S. Lamberth, A. Bester-van der Merwe, P. Cowley, P. Bloomer, and R. Roodt-Wilding. 2016. Genetic analyses reveal declining trends and low effective population size in an overfished South African sciaenid species, the dusky kob (Argyrosomus japonicus). Marine and Freshwater Research 67(2):266.

Nunney, L., and D. R. Elam. 1994. Estimating the effective population size of conserved populations. Conservation biology 8(1):175–184. Wiley Online Library.

Ottewell, K. M., D. C. Bickerton, M. Byrne, and A. J. Lowe. 2016. Bridging the gap: a genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. Diversity and Distributions 22(2):174–188.

Palstra, F. P., and D. J. Fraser. 2012. Effective/census population size ratio estimation: a compendium and appraisal. Ecology and Evolution 2(9):2357–2365.

Palstra, F. P., and D. E. Ruzzante. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Molecular Ecology 17(15):3428–3447.

Ralls, K., J. D. Ballou, M. R. Dudash, M. D. B. Eldridge, C. B. Fenster, R. C. Lacy, P. Sunnucks, and R. Frankham. 2018. Call for a Paradigm Shift in the Genetic Management of Fragmented Populations: Genetic management. Conservation Letters 11(2):e12412.

Rosser, A. M., and S. A. Mainka. 2002. Overexploitation and Species Extinctions. Conservation Biology 16(3):584–586.

Ruzzante, D. E., G. R. McCracken, S. Parmelee, K. Hill, A. Corrigan, J. MacMillan, and S. J. Walde. 2016. Effective number of breeders, effective population size and their relationship with census size in an iteroparous species, *Salvelinus fontinalis*. Proceedings of the Royal Society B: Biological Sciences 283(1823):20152601.

Therkildsen, N. O., E. E. Nielsen, D. P. Swain, and J. S. Pedersen. 2010. Large effective

population size and temporal genetic stability in Atlantic cod (Gadus morhua) in the southern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences 67(10):1585–1595.

Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. Philosophical Transactions of the Royal Society B: Biological Sciences 360(1459):1395–1409. The Royal Society London.

Wang, J. 2016. A comparison of single-sample estimators of effective population sizes from genetic marker data. Molecular Ecology 25(19):4692–4711. Wiley Online Library.

Waples, R. S. 2002. Definition and estimation of effective population size in the conservation of endangered species. Population viability analysis:147–168. University of Chicago Press.

Wilson, M. C., X.-Y. Chen, R. T. Corlett, R. K. Didham, P. Ding, R. D. Holt, M. Holyoak, G. Hu, A. C. Hughes, L. Jiang, W. F. Laurance, J. Liu, S. L. Pimm, S. K. Robinson, S. E. Russo, X. Si, D. S. Wilcove, J. Wu, and M. Yu. 2016. Habitat fragmentation and biodiversity conservation: key findings and future challenges. Landscape Ecology 31(2):219–227.

**Appendix A: Molecular Protocol**

Forward and reverse primers for each one of the 33 loci were purchased from Integrated DNA technologies (IDT, Coralville, IA, USA) tailed with Illumina (San Diego CA, USA) Read1_(CCCTACACGACGCTCTTCCGATCT) and Read2_(GTTCAGACGTGTGCTCTTCCGATCT) sequencing primers, respectively. A single multiplex PCR consisting of all 33 loci was performed per individual using the Qiagen Multiplex PCR Kit (Qiagen Inc., Valencia, CA, USA) following the manufacturers recommendations, scaled back to 5µl total reaction volume. The multiplex PCRs were carried out on Eppendorf (Hamburg, Germany) Mastercycler ep384 PCR machines with the following parameters: 94°C for 15m, 20 cycles of 94°C for 30s, 57°C for 180s, 72°C for 60s, and a final extension at 68°C for 30 minutes.

The resulting multiplex PCR products were diluted with 20µl of purified water in preparation for indexing PCR (see Zhan *et al.,* 2017 for details). Indexing PCRs were carried in 5µl reaction volumes consisting of 1.95µl of purified water, 0.5µl of 10x buffer, 0.2 mM of each dNTP, 0.2µM of each index oligo, 0.5µl of diluted multiplex PCR product, and 0.25U of TSG (Bio Basic, Markham, ON, Canada). The following thermocycler parameters were used: 95°C for 120s, 20 cycles of 95°C for 20s, 60°C for 60s, 72°C for 60s, and a 72°C final extension for 10 minutes.

PCR products were then pooled in equal proportion and cleaned using a 1.8:1 ratio of Sera-Mag Speedbeads (GE Healthcare, Little Chalfront, UK) to pooled PCR library. Library quantification was completed on a Roche LC480 qPCR machine (Roche, Basel, Switzerland) using the appropriate Kapa Library Quantification Kit (*Roche, Pleasanton, California)* for use with the Illumina sequencing platform. We followed the manufacturer's guidelines for this quantification. Libraries were subsequently diluted to 15pM and sequenced in a single direction using an Illumina MiSeq Benchtop Sequencer (Illumina, San Diego, CA, USA). Due to expected

fragment size, we utilized MiSeq 150 cycle V3 chemistry kits with dual indexing, and sequenced 150bp in one direction, rather than paired end sequencing.

Post sequencing, dual indexed individuals were demultiplexed automatically using the MiSeq Sequence Analysis software. This resulted in the output of a single FASTQ file per individual, containing sequence data for all microsatellite loci related to that individual. This information was then input to MEGASAT (Zhan *et al.,* 2017) which further demultiplexed individuals sequence data, based on loci, using locus specific information. Simultaneously, this software genotyped all loci, and output depth histograms for manual verification of scoring accuracy/consistency.

**Table A1**: Reference information on 33 loci used in microsatellite analyses.

| Locus name | Reference | No tail left | Rev. comp of no tail right | Five Flank | Three Flank | Repeat |
|---|---|---|---|---|---|---|
| SFOC88 | King et al. (2012)[1] | GGGAGAACCCAGTGTTTCTTT | CGTTCACAATCAGGGTTCAG | ATTTGACCTAAGTGAACTGTGTTATAAG | GA | GAT |
| SFOC24 | King et al. (2012)[1] | AACACTGGAGCCGTTGAAGT | TTAGGCATCACCCCATCTCT | TATGGGT | G | GAT |
| SFOD129 | King et al. (2012)[1] | GTGCAGGCACTAACTGGACA | TGAAGATGAGGATTCCCTGG | GAGCCTGAGGACCCTGAGGATGAAGATGACGTTCCT | GAAGAGGACCCTCTTATATT | GAT |
| SFOC28 | King et al. (2012)[1] | CAGTTGAAGTGATTGGGTTAGC | GCTTTTGTGTGTGTGGTGTG | TACTTCTGTTTCTTTCTCTGTGTGTGTGT | GC | GCGT |
| SFOC113 | King et al. (2012)[1] | GGGGAGCCCAGACTATATTGA | TACCCTGATGGCAATGATGA | CGATGACAACTACATAGACAGT | GAA | GAT |
| Ssa-1.14 | Bradbury et al. (2018)[2] | TCGTATTTGTCAAGGATGTGCC | GGGCAATACAATGGGCATCT | TTG | TTTATT | AGT |
| Ssa-1.7 | Bradbury et al. (2018)[2] | AGAACACAACAGAACCAGGTAC | GGGTTGGAAGTGTGTTCGAG | TACTGACCATGGTCCT | GGTCAATAATCCCTCACACTG | GAT |
| Ssa-10.2 | Bradbury et al. (2018)[2] | TGATCCTCTTCACCACCCTG | GGTGAGGGAGGAGTCTTCAG | TGTTCTCCACAT | ATCAT | AAT |
| Ssa-15.1 | Bradbury et al. (2018)[2] | TTTCTTTGTGTGTTGTGCCC | CCCAGAGGAACCACAGCTG | TTTAATCATC | CAGTTGTCCCTGGC | CCT |
| Ssa-21.5 | Bradbury et al. (2018)[2] | CACTCCCTAACTCCATGGTC | CACAGTGACGACATCCATGA | ATCACATGGTCCTAGTACTGAACGGGAAAG | AAAGGACAAACGCAAATAGAGGACGT | GAT |
| Ssa-26.d06 | Bradbury et al. (2018)[2] | CATAATCACCTTGCATGACACC | GTATTTAGCGGTGCAGCAGG | TTTAT | CT | AC |
| Ssa-01.12 | Ian Paterson, unpublished | ACAATGTCGCT | GTAATCGGCGGCTACAGATG | GCTGTACT | GGATCTCCATGTGGTC | GCT |
| Ssa-05.10 | Ian Paterson, unpublished | GCTTCCACGCCCATAACAAC | GAGAGCCGAGGAGTTGGAG | AACATGGC | GGTGGAGAATAACAAC | AGG |
| Ssa-03.7 | Lien et al. (2016)[3] | GCACATTGAAGTTGGTTGCC | GACGAGCAGCTTCTGGTTAA | AACCACCATATAAAAACCA | X | AAG |

| | | | | | | |
|---|---|---|---|---|---|---|
| Ssa-04.d56 | Lien et al. (2016)[3] | CTGCTGGTAAATGGGCGTTG | GCCTTTGTTAACCAGGACAGG | GTGCA | TTTGAA | GT |
| Ssa-06.8 | Lien et al. (2016)[3] | TGAGGCCGATGTCACCTG | GCCGTCGAGTTCAGGTACT | GATGTGACCTGTCACCGTCATC | CCCTGTTCGGAAT | CCT |
| Ssa-09.12 | Lien et al. (2016)[3] | CTGACAGGTGGAGTGGGAC | CTGCAGGTACATGCGGGA | CCCTCTCCATC | GTTCGTCTGG | TCC |
| Ssa-10.3 | Lien et al. (2016)[3] | TGATGGGTCTTGGTGTAGGG | CCTGGGCTTCACCGTTGA | GTGATAGA | GGTGGTCTCTCCCTGGCTCTGGC | AGT |
| Ssa-11.1 | Lien et al. (2016)[3] | AGAGCTCCGACACACATTCG | GCACCGGCCTAGCTCTATG | GCCG | CAGGCAGCACTGCATTTTTGA | CCT |
| Ssa-12.2 | Lien et al. (2016)[3] | ACTGGTAGGTCATTGTTCTGTG | GCACCGAGAACACACATAAGG | TTTTGATCAGCACACATTATTTTTTATTTTATTACTATT | CT | ACT |
| Ssa-13.6 | Lien et al. (2016)[3] | GCTGTTCCTCTGGCCTCAC | GATAGTACAGTGTTGAGGTGCT | CTCCTTGT | TAGTCTGGTCA | TCC |
| Ssa-14.10 | Lien et al. (2016)[3] | GGGAACGTGTGGAAGATTCAC | GGCATCACCCTCCATACCTT | TCACTAAGGTCCTC | AGC | ATC |
| Ssa-15.9 | Lien et al. (2016)[3] | ATACTACCTGTTCAGGCGGC | GGGAGGAGAGTCATCACAGG | CACCGG | A | AGG |
| Ssa-16.2 | Lien et al. (2016)[3] | GTTTACGTCACCTGCAGCTG | CGAGGGCTTAACATCTACTGC | AGCCCCCTCCACTCT | ACCACCACTA | ACT |
| Ssa-20.3 | Lien et al. (2016)[3] | GGAGGGAGTGTAGAGGCTTTC | GTAGCAGAGATGGGTGTGTG | AGTTGAATTGTTTACCACG | T | AGG |
| Ssa-20.d16 | Lien et al. (2016)[3] | GGCAACGAGGTGAGAATGC | GGCAACTAGGTAAGACGCAC | GC | GGTCA | AC |
| Ssa-23.9 | Lien et al. (2016)[3] | ACGGATACAGAGAGACGCAC | GACTTTGTCCTCCTCGCTGT | GCGGGATATACTAGTGTTATCAGTATTTTCTTT | AAGGTGGTGGTATCACCTTTAG | ATC |
| Ssa-27.1 | Lien et al. (2016)[3] | TCCATGAGTACACGCCACTG | CCAGGGTAGAGTAGTGGAGAAC | ACT | X | CT |
| Ssa-27.d07 | Lien et al. (2016)[3] | GATTTCACAAAGCAGCGCG | CTCGCGACAGAACATGCTG | CGCCAGTGTACC | GCGGGTC | AC |
| Ssa-27.d19 | Lien et al. (2016)[3] | GGAATACTGTCTCATTGCGCC | CACCTACAACCTTTGATTGCCT | TCAAGTGTGACT | CGAGAGATTA | GT |

| | | | | | | |
|---|---|---|---|---|---|---|
| Ssa-28.d08 | Lien et al. (2016)[3] | TCTGACCTACACAC AACAATGG | CATTCTGAGCGAG CACACAC | AAGAGCTA | AAACA | AC |
| Ssa-29.2 | Lien et al. (2016)[3] | GGCACAGCACACCA GTTG | GGAACATCTTGGA ACGCTGT | TCTCTCCTTCTCTTCT | CACTCCTCTCATCT GGTCTTCTTTA | TCC |
| Ssa-4.9 | Lien et al. (2016)[3] | AGAATCTCTAGCCC ACACAAC | GCTCACATCTCAA CCCTGC | X | TA | AAC |

[1]King, T. L., Lubinski, B. A., Burnham-Curtis, M. K., Stott, W., & Morgan, R. P. (2012). Tools for the management and conservation of genetic diversity in brook trout (Salvelinus fontinalis): tri-and tetranucleotide microsatellite markers for the assessment of genetic diversity, phylogeography, and historical demographics. *Conservation Genetics Resources*, *4*(3), 539-543.

[2]Bradbury, I. R., Wringe, B. F., Watson, B., Paterson, I., Horne, J., Beiko, R., ... & Bentzen, P. (2018). Genotyping-by-sequencing of genome-wide microsatellite loci reveals fine-scale harvest composition in a coastal Atlantic salmon fishery. *Evolutionary Applications*, *11*(6), 918-930.

[3]Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., ... & Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*(7602), 200-205.

**Appendix B: Conversion of $N_b$ to $N_e$ Using Two Life-History Traits**

We calculated $\widehat{N}_{e(adj)}$ from $\widehat{N}_b$ using age-at-maturity ($\alpha$) and adult lifespan (AL) according to the formulae using two life-history traits in Waples et al. (2014). Age-at-maturity in our populations was calculated based on fish captured during stock assessment, using generalized linear mixed models with a binomial fit to estimate age-at-fifty-percent-maturity ($A_{50}$, here we refer to as $\alpha$). Adult lifespan was approximated by taking the age of the oldest fish caught in each lake per year (Table B1). To calculate $\widehat{N}_{e(adj)}$, we first took the raw $\widehat{N}_b$ values and calculated adjusted $\widehat{N}_b$ (or $\widehat{N}_{b(adj)}$) as follows:

$$\widehat{N}_{b(adj)} = \frac{raw\ \widehat{N}_b}{1.103 - 0.245 \times \log\left(\frac{AL}{\alpha}\right)}$$

Then, using $\widehat{N}_{b(adj)}$, we calculated $\widehat{N}_{e(adj)}$:

$$\widehat{N}_{e(adj)} = \frac{\widehat{N}_{b(adj)}}{0.485 + 0.758 \times \log\left(\frac{AL}{\alpha}\right)}$$

We also calculated the ratio between $\widehat{N}_{e(adj)}$ and $\widehat{N}_c$ to assess whether genetic compensation was detected with the converted effective sizes.

**Results**

$\widehat{N}_{e(adj)}$ followed the same trends through time as $\widehat{N}_b$, with only Mud having a significant change in $\widehat{N}_{e(adj)}$ through time (indicated by non-overlapping confidence intervals; Figure B1). The ratio between $\widehat{N}_{e(adj)}$ and $\widehat{N}_c$ also had the same trends as with $\widehat{N}_b/\widehat{N}_c$, with all the harvest lakes increasing in the ratio through time, indicating that genetic compensation has occurred.

**Table B1**: Effective population size ($N_{e(adj)}$) calculated from the effective number of breeders ($N_b$) using age-at-maturity ($\alpha$) and adult lifespan (AL). $N_{b(adj)}$ is the adjusted $N_b$ based on the two life-history traits.

| | Population - Year | $\alpha$ | AL | $\widehat{N}_{b(adj)}$ | $\widehat{N}_{e(adj)}$ | $\widehat{N}_{e(adj)}/\widehat{N}_c$ |
|---|---|---|---|---|---|---|
| Control | Helen 2017 | 1.17 | 5 | 70.2 | 72.9 | 0.08 |
| | Helen 2019 | 1.13 | 5 | 122.4 | 125.6 | 0.28 |
| | Helen 2020 | 2.14 | 6 | 69.5 | 84.3 | 0.12 |
| | Marg 2017 | 1.57 | 4 | 87.5 | 110.4 | 0.07 |
| | Marg 2019 | 1.56 | 5 | 45.5 | 52.3 | 0.03 |
| | Marg 2020 | 1.87 | 5 | 73.9 | 91.4 | 0.05 |
| | McNair 2017 | 2.32 | 8 | 46.2 | 51.8 | 0.21 |
| | McNair 2019 | 2.29 | 6 | 17.9 | 22.3 | 0.13 |
| | McNair 2020 | 3.1 | 7 | 22.6 | 30.0 | 0.14 |
| Harvest | Mud 2017 | 1.38 | 6 | 61.8 | 63.8 | 0.10 |
| | Mud 2019 | 0.77 | 5 | 26.2 | 23.8 | 0.07 |
| | Mud 2020 | 1 | 3 | 21.8 | 25.8 | 0.28 |
| | Olive 2017 | 2.04 | 5 | 50.9 | 65.3 | 0.03 |
| | Olive 2019 | 1.46 | 4 | 34.8 | 42.7 | 0.05 |
| | Olive 2020 | 1.68 | 4 | 54.4 | 70.6 | 0.22 |
| | Temple 2017 | 3.29 | 7 | 26.8 | 36.5 | 0.04 |
| | Temple 2019 | 2.44 | 8 | 69.5 | 79.4 | 0.21 |
| | Temple 2020 | 2.76 | 10 | 45.8 | 50.3 | 0.26 |

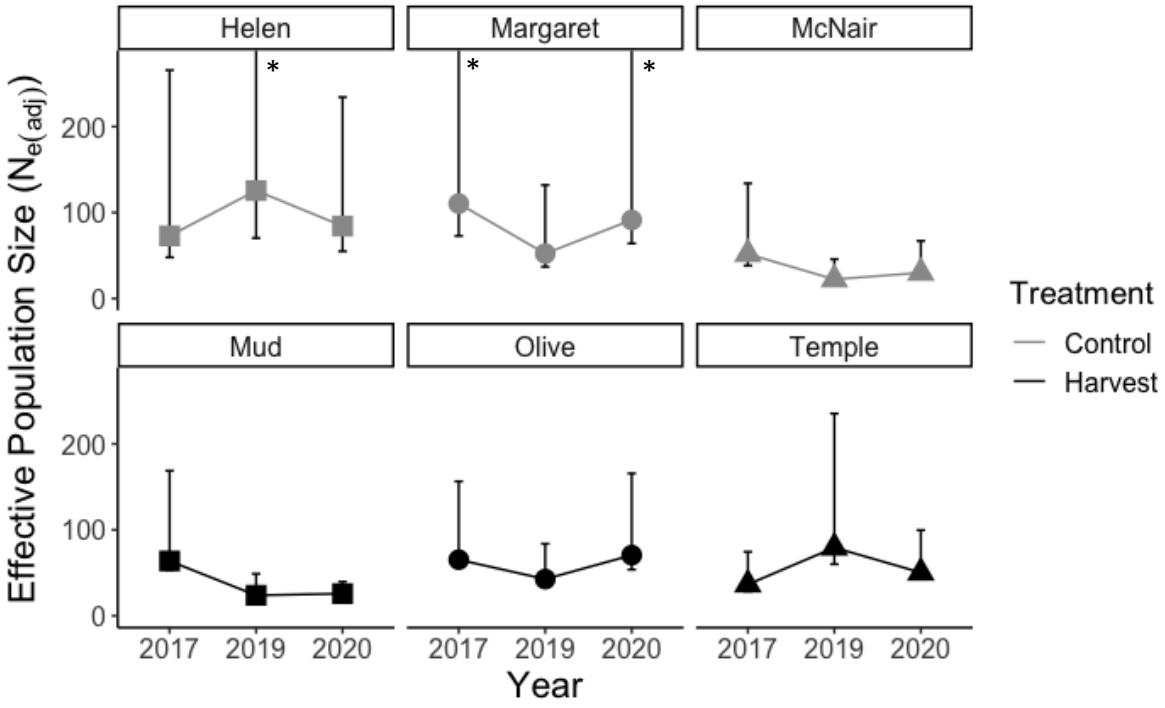


**Figure B1:** Effective population size ($N_{e(adj)}$) through time in control and harvest populations of brook trout in the Rocky Mountains. Error bars indicate converted 95% confidence intervals. Asterisks indicate where the error bar did not fit within the limits of the graph.

**Appendix C: Supplementary Material**

**Table S1:** Environmental characteristics of six alpine lakes inhabited by brook trout populations in the Rocky Mountains

| Lake | Treatment | Species* | Surface area (ha.) | Elevation (m) | Maximum Depth (m) |
|---|---|---|---|---|---|
| Helen | Control | 1 | 2.48 | 2400 | 15 |
| Margaret | Control | 1,3 | 18 | 1808 | 28.2 |
| McNair | Control | 1 | 1.7 | 1532 | 4 |
| Mud | Harvest | 1,2 | 7.2 | 1600 | 7.2 |
| Olive | Harvest | 1 | 1.66 | 1470 | 3.6 |
| Temple | Harvest | 1 | 3.25 | 2207 | 14.7 |

*Species – 1: Brook Trout, 2: Longnose Dace, 3: Westslope Cutthroat Trout

**Table S2**: Number of fish harvested and harvest rates (in brackets) across experimentally harvested populations of brook trout in the Rocky Mountains

| | Lake | 2017 | 2018 | 2019 |
|---|---|---|---|---|
| Harvest | Mud | 521 (0.45) | 478 (0.59) | 378 (0.83) |
| | Olive | 758 (0.27) | 811 (0.46) | 608 (0.65) |
| | Temple | 1172 (0.58) | 1161 (0.75) | 703 (0.78) |

**Table S3**: Census size ($N_c$) and 95% CIs over three successive years for control and harvest alpine brook trout populations

| | Lake | 2017 | | | 2018 | | | 2019 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $N_c$ | LCI | UCI | $N_c$ | LCI | UCI | $N_c$ | LCI | UCI |
| Control | Helen | 988 | 682 | 1484 | 504 | 385 | 729 | 777 | 553 | 1127 |
| | Margaret | 1684 | 1129 | 2619 | 1923 | 1569 | 2483 | 2052 | 1569 | 2967 |
| | McNair | 269 | 152 | 510 | 195 | 153 | 266 | 245 | 183 | 335 |
| Harvest | Mud | 1150 | 883 | 1648 | 805 | 690 | 968 | 456 | 387 | 557 |
| | Olive | 2800 | 2038 | 3953 | 1747 | 1369 | 2413 | 934 | 759 | 1215 |
| | Temple | 2012 | 1538 | 2906 | 1533 | 1277 | 1918 | 896 | 673 | 1218 |

**Table S4**: Expected and observed heterozygosity ($H_e$, $H_o$), and number of alleles ($N_a$) for control and harvest populations of brook trout in the Rocky Mountains. $N_{micro}$ indicates the sample size used in the microsatellite analyses

| Population - Year | $N_{micro}$ | $H_e$ | $H_o$ | $N_a$ |
|---|---|---|---|---|
| Helen 2017 | 61 | 0.435 | 0.453 | 3.214 |
| Helen 2019 | 65 | 0.423 | 0.416 | 3.321 |
| Helen 2020 | 65 | 0.440 | 0.445 | 3.393 |
| Marg 2017 | 79 | 0.360 | 0.368 | 2.870 |
| Marg 2019 | 90 | 0.376 | 0.38 | 2.364 |
| Marg 2020 | 88 | 0.367 | 0.364 | 2.409 |
| McNair 2017 | 43 | 0.443 | 0.459 | 3.037 |
| McNair 2019 | 46 | 0.406 | 0.407 | 3.037 |
| McNair 2020 | 43 | 0.420 | 0.425 | 2.926 |
| Mud 2017 | 68 | 0.472 | 0.467 | 3.074 |
| Mud 2019 | 72 | 0.458 | 0.478 | 2.926 |
| Mud 2020 | 74 | 0.479 | 0.485 | 2.963 |
| Olive 2017 | 74 | 0.41 | 0.390 | 3.111 |
| Olive 2019 | 76 | 0.408 | 0.402 | 3.286 |
| Olive 2020 | 72 | 0.417 | 0.402 | 3.346 |
| Temple 2017 | 66 | 0.400 | 0.411 | 3.037 |
| Temple 2019 | 74 | 0.441 | 0.451 | 3.357 |
| Temple 2020 | 74 | 0.449 | 0.458 | 3.429 |
| Mean | 68.3 | 0.422 | 0.426 | 3.061 |

**Table S5**: Effective number of breeders ($N_b$) and jackknife 95% CIs for control and harvest alpine brook trout populations

| | Lake | 2017 | | | 2019 | | | 2020 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $N_b$ | LCI | UCI | $N_b$ | LCI | UCI | $N_b$ | LCI | UCI |
| Control | Helen | 66.6 | 35.9 | 199.3 | 115.6 | 55.6 | 746.5 | 69.0 | 39.2 | 167.3 |
| | Margaret | 87.8 | 45.7 | 279.8 | 44.5 | 25.3 | 90.5 | 73.8 | 39.3 | 197.5 |
| | McNair | 44.9 | 26.9 | 94.3 | 17.9 | 11.2 | 29.6 | 23.0 | 14.3 | 40.2 |
| Harvest | Mud | 58.5 | 36.0 | 113.8 | 23.7 | 16.2 | 35.3 | 21.5 | 16.0 | 29.0 |
| | Olive | 51.3 | 33.5 | 88.1 | 34.7 | 24.1 | 52.5 | 55.0 | 34.0 | 104.8 |
| | Temple | 27.4 | 17.4 | 45.6 | 67.9 | 39.2 | 153.6 | 44.2 | 30.8 | 67.5 |

**Table S6**: Ratio between effective and census size ($N_b/N_c$) in control and harvest alpine brook trout populations

| | Lake | 2017 | 2019 | 2020 |
|---|---|---|---|---|
| Control | Helen | 0.067 | 0.255 | 0.098 |
| | Margaret | 0.052 | 0.025 | 0.038 |
| | McNair | 0.167 | 0.107 | 0.111 |
| Harvest | Mud | 0.051 | 0.072 | 0.234 |
| | Olive | 0.018 | 0.037 | 0.169 |
| | Temple | 0.014 | 0.181 | 0.226 |

**Table S7**: Backwards model selection for Generalized Linear Mixed Models comparing the influence of census size and life history traits on $N_b$ in control and harvest alpine brook trout populations.

| Model | Model Terms | Log Lik. | AIC | ΔAIC |
|---|---|---|---|---|
| Census Size | $N_b \sim 1 + (1 \mid Lake)$ | – 80.76 | 169.23 | 0.00 |
| | $N_b \sim Treatment + (1 \mid Lake)$ | – 80.57 | 172.22 | 3.00 |
| | $N_b \sim N_c + (1 \mid Lake)$ | – 80.73 | 172.54 | 3.32 |
| | $N_b \sim N_c + Treatment + (1 \mid Lake)$ | – 80.42 | 175.85 | 6.63 |
| | $N_b \sim N_c * Treatment + (1 \mid Lake)$ | – 80.41 | 180.46 | 11.23 |
| CV Body Length | $N_b \sim 1 + (1 \mid Lake)$ | – 80.76 | 169.23 | 0.00 |
| | $N_b \sim CV\ length + (1 \mid Lake)$ | – 80.31 | 171.70 | 2.47 |
| | $N_b \sim Treatment + (1 \mid Lake)$ | – 80.57 | 172.22 | 3.00 |
| | $N_b \sim CV\ length + Treatment + (1 \mid Lake)$ | – 79.97 | 174.94 | 5.72 |
| | $N_b \sim CV\ length * Treatment + (1 \mid Lake)$ | – 79.96 | 179.56 | 10.33 |
| Sex Ratio | $N_b \sim 1 + (1 \mid Lake)$ | – 80.76 | 169.23 | 0.00 |
| | $N_b \sim Sex\ Ratio + (1 \mid Lake)$ | – 80.51 | 172.09 | 2.87 |
| | $N_b \sim Treatment + (1 \mid Lake)$ | – 80.57 | 172.22 | 3.00 |
| | $N_b \sim Sex\ Ratio + Treatment + (1 \mid Lake)$ | – 80.34 | 175.68 | 6.45 |
| | $N_b \sim Sex\ Ratio * Treatment + (1 \mid Lake)$ | – 80.05 | 179.74 | 10.52 |

**Table S8** – Single-sample genetic $N_e$ estimation methods included in screening criteria

| Estimation Method | Description | Software | Reference |
|---|---|---|---|
| Approximate Bayesian Computation | Uses summary statistics from a population with Bayesian computation to provide posterior probability distributions for $N_e$. | ONeSAMP | (Tallmon et al., 2008) |
| Heterozygote Excess | Estimates the effective number of breeders based on the assumption that when $N_b$ is small, there will be an excess of heterozygotes (Pudovkin et al., 1996; Zhdanova & Pudovkin, 2008). | NeEstimator v2 | (Do et al., 2014) |
| Linkage Disequilibrium (LD) | Uses non-random associations of alleles at different loci (i.e. deviations from the expected genotype frequency based on random distribution) to estimate random drift and therefore $N_e$ or $N_b$ (W. G. Hill, 1981). A bias correction by Waples (2006) eliminates downward bias when sample size is less than the true $N_e$. | LDNe<br>NeEstimator v2 | (Waples & Do, 2008)<br>(Do et al., 2014) |
| Molecular Coancestry | Estimates the effective number of breeders from molecular coancestry, a measure of shared alleles between individuals (Nomura, 2008). | NeEstimator v2 | (Do et al., 2014) |
| Sibship Frequency | Estimates demographic parameters from genotypes of a cohort, and then uses these with predictive equations based on the probabilities of sampling half-sibs or full-sibs from the population (Wang, 2009). | COLONY | (Jones & Wang, 2010) |

**Table S9** – Protocol used by reviewers to extract data from articles included in the $N_e$ database

| Group | Column | Explanation |
|---|---|---|
| | Common Name | name used by authors in article, all lower case |
| | Genus | genus (capitalized) |
| | Species | species (uncapitalized) |
| | Taxonomic group | freshwater fish, marine fish, anadromous fish, reptile, amphibian, mammal, bird, invertebrate, plant. If something is not listed here, or you are unsure, you can enter a comment in the column next to this. For fish that can be either resident or anadromous, use best judgement according to the authors' descriptions of the population |
| | Population | name given to population; for species in bodies of water, could use the name of the lake/river, or another name used by authors in article |
| | Population ID | numerical identification for each population (since there can be multiple estimates for a single population). Alpha-numeric system using the article number. i.e. if article # 100 has two populations, they will be 100A and 100B. If an article only has one estimate, still include A at the end. |
| **Population information** | Method of defining population | based off of the authors in the article and how they defined the population. E.g. using Fst values, STRUCTURE (determining # of groups), BAYESASS (measuring migration rates), IBA (individual-based-assigment; using genetic data from populations to assign individuals), etc. If there is any additional information, include it in the comment column. E.g. what their threshold Fst value was, or level of migration, etc. |
| | Location (coordinates) | coordinates in lat/long, decimal degrees, to 6 decimals if available. If article has info on UTM, use this website to convert: https://www.engineeringtoolbox.com/utm-latitude-longitude-d_1370.html . If article has info on lat/long in degrees/minutes/seconds, use this website to convert: https://www.latlong.net/degrees-minutes-seconds-to-decimal-degrees . If article includes a map or description of location but does not provide coordinates, use google maps to generate coordinates as accurately as possible (e.g. if given the name of a lake, take the coordinates in middle of lake). You can also use google maps to find the "average" coordinates if you are pooling sampling locations that make up a single population (helpful tool is "measure distance" and can take the midpoint) |
| | Region where population is located | can be a city/province/ etc. or multiple of these things. |

| | | |
|---|---|---|
| | Country | free-form text. Please capitalize country name, and use as accurate spelling as possible. If samples were taken from multiple countries, you can enter one here, and add the others in the "region" column. |
| | Continent/Ocean | drop-down list of contintents and Oceans. For anything not included in the drop-down list, you can enter it in the "region" column. E.g. |
| | Was the pop stocked or supplemented? | e.g. historical supplementation, stocking into a new location, re-introductions (that are now reproducing independently) |
| | Is the pop non-native? | non-native or invasive species |
| | When was the pop stocked/ introduced? | the year when pop was introduced or last stocked |
| | notes | notes on population information |
| **Protection status** | At-risk | YES/NO (based on information in the article) |
| | Protection status | threatened, vulnerable, endangered, etc. (enter "other" into comments column). Based on what is written in article by authors |
| | Protection status organization | e.g. IUCN, COSEWIC, USFW, etc. (population-specific info is better if mentioned, instead of IUCN) |
| **Method** | Method (general) | LD (linkage disequilibium), SF (sibship frequency), HE (heterozygote excess), MC (molecular coancestry), Bayesian methods. If article used a different single-sample genetic estimator not listed, put "Other" and enter it into the notes column |
| | Method (specific) | LDNe, NeEstimator (v1 or V2), COLONY, ONeSAMP. If article used a different software not listed, put "Other" and enter it into the notes column |
| | Marker type | microsatellite, allozyme, SNP, etc. |
| | GW correction | genome-wide bias correction for LD method; YES/NO. Based on study by waples |
| | Allele freq cutoff | for LD method; common values are 0.01, 0.05, 0.1. If the article reports multiple allele cutoff values, follow this rule: For sample sizes >25 use 0.02, and <25 use 0.05. Please report whether the study included multiple allele cutoffs. |

| | | |
|---|---|---|
| | Mating system | for SF method. If the mating system is not one of the given options, please choose "other" and enter it in the comments column. If the study includes data on multiple mating systems, choose the best option based on the authors' discussion, and make a comment that there were multiple values reported |
| | other comments for SF method | here, please include any other details about the SF method that the authors report. E.g. whether it was based on random mating or inbreeding; the probability of an offspring in a dataset having a parent in the dataset; etc. If multiple values are given for any of these, choose one either based on author discussion, or choose the median value (e.g. if chance of sampling a parent is reported at 0.3, 0.5 and 0.7, use 0.5), and make a comment that other values are reported in the article |
| | Priors | For the Bayesian method in ONeSAMP. Upper and lower boundaries of the Ne made prior to estimation. (e.g. 5-150). If multiple values are given, choose the estimate with the widest range of priors. (i.e. choose 2-50 rather than 10-20, etc.) |
| | Type of sequencing | RAD-seq, GBS, capillary electrophoresis, Sanger. Can enter a method not listed here by choosing "other" and then entering in comment column. |
| | # of loci/SNP | number of loci or SNP used in estimate. **ensure you adjust this number if the study excludes monomorphic loci or loci with high null freq.** |
| | He and Ho | average He (expected heterozygosity) or Ho (observed heterozygosity) across loci for that population. For microsatellites. |
| | Ar | allelic richness; average # of alleles per locus, weighted by sample size. If the authors refer to "Ar" with no mention of method, assume they are correct. If they refer to Ar as the non-weighted version, then please categorize as MNA instead. |
| | MNA | mean number of alleles per locus. NOT weighted. If the authors refer to MNA but mention weighting, categorize as Ar instead. If they refer to MNA with no mention of method, assume they are correct. |
| | Inbreeding coefficient (Fis) | usually calculated from heterozygosity measures. |
| | Nucleotide diversity | for SNPs |
| | Sample size | the number of individuals sampled from the population |
| **Ne values** | Ne | point estimate of Ne. If this comes from several sample sites for same pop (e.g. Durrant et al 2011), remove the population/study but flag it. If the value is infinite, enter "infinity" (please record these for now; though they may be excluded later) |

| | | |
|---|---|---|
| | Nb | point estimate of Nb |
| | Year estimate was taken | year the samples were taken from the population. If no year is given, leave blank. If multiple years, use most recent year but make note of other years. |
| | LCI | lower confidence interval for estimate (can be "infinity" as well) |
| | UCI | upper confidence interval for estimate (can be "infinity" as well) |
| | CI method | method used to calculate Cis. E.g. jackknife vs parametric methods in LDNe program. If the method is not provided, enter as text in comments column. |
| **Ne verification** | Did they sample across cohorts? | yes/no/unsure. in order to be an accurate estimate of Ne, there need to be sampling across different cohorts (i.e. birth years). If only one cohort is sampled, then this is Nb, not Ne. |
| | Did they report Ne for sampling sites? | If the authors define a population as a group of sampling sites, but only report Ne for the sites (rather than for the population as a whole), please mark this column as "YES", and report each sampling site on a unique row, with the SAME POPULATION ID. If Ne is reported for both the sampling sites, and overall population, use the population-level Ne, and mark this column as NO, because you are reporting on the population level. |
| | Did they pool samples from multiple years? | yes/no/unsure . Put other years in the notes column |
| | Notes | any other notes on the validity of the Ne estimate (e.g. all samples came from a single breeding site and may not be representative of the population) |
| **Nc values** | Nc estimate and Cis | same as Ne, but with Nc |
| | Nc method | mark-recapture (in comment, please put the type of mark-recapture. i.e. petersen/lincoln-petersen, schnabel, etc.), complete count, incomplete count (e.g. quadrat study with extrapolation). If the method is not in the list, please put "other" and then write out method in comment column. |

**Table S10** – Replication accounted for within the $N_e$ review database

| Type of Replication | Description | Which records were retained? |
|---|---|---|
| Marker Replication | For a single population, the authors used both microsatellites and SNPs to estimate $N_e$ or $N_b$. | SNP estimates, as they had a higher power (which we calculated by multiplying the number of markers, sample size, and allelic richness) |
| Software Replication | Authors estimated $N_e$ or $N_b$ using both NeEstimator V2 and LDNe | Estimates generated in NeEstimator V2, as it is the more recent software, and includes an updated bias correction to address missing data (Do et al., 2014) |
| Spatial Replication | Multiple estimates from a single population (i.e. they all shared the same population ID), for example when $N_e$ was reported for several sampling sites that were shown through Fst or STRUCTURE to cluster into one population | The estimate with the largest sample size (S), as larger sample sizes generally produce more reliable estimates (Waples & Do, 2010). In the case where two estimates with the same population ID had the same sample size, we chose either the estimate with a non-infinite upper confidence interval, or we chose randomly (using a random number generator) if all upper CIs were finite. |
| Temporal Replication | Multiple estimates were reported over time for the same population | The most recent estimate, as these were most relevant for a contemporary review of $N_e$ |

**Table S11** – Sample sizes by taxonomic group for all Generalized Linear Mixed Models of $N_e$ database

| Taxonomic Group | $N_e$ Analysis | $N_b$ Analysis | 50/500 Analyses | Human Footprint Analysis | Ratio Analysis |
|---|---|---|---|---|---|
| Amphibians | 369 | 68 | 369 | 339 | 25 |
| Anadromous Fishes | 766 | 210 | 766 | 586 | 107 |
| Birds | 204 | 9 | 204 | 165 | 41 |
| Freshwater Fishes | 1033 | 357 | 1033 | 965 | 140 |
| Invertebrates | 333 | 60 | 333 | 226 | 6 |
| Mammals | 552 | 69 | 552 | 477 | 96 |
| Marine Fishes | 248 | 39 | 248 | 22 | 9 |
| Plants | 439 | 5 | 439 | 387 | 80 |
| Reptiles | 201 | 10 | 201 | 132 | 33 |
| **Total** | **4145** | **827** | **4145** | **3299** | **537** |

**Table S12** – Backward selection of Generalized Linear Mixed Models for the $N_e$ review using AICc. Bold indicates the most parsimonious model, and * indicates interactions

| Analysis | Equation* | df | AICc | ΔAICc |
|---|---|---|---|---|
| $N_e$ | $N_e$ ~ taxonomic group + marker + loci number + (1 \| StudyID / PopulationID) | 15 | 3121919 | 0.000 |
| | **$N_e$ ~ taxonomic group + marker + (1 \| StudyID / PopulationID)** | 14 | 3121917 | –1.724 |
| $N_b$ | $N_b$ ~ taxonomic group + marker + loci number + (1 \| StudyID / PopulationID) | 14 | 970312.2 | 0.000 |
| | **$N_b$ ~ taxonomic group + marker + (1 \| StudyID / PopulationID)** | 13 | 970310.7 | –1.452 |
| Ratio | Ratio ~ taxonomic group + marker + loci number + ratio type + (1 \| StudyID / PopulationID) | 15 | –102684.2 | 0.000 |
| | Ratio ~ taxonomic group + loci number + ratio type + (1 \| StudyID / PopulationID) | 14 | –102686.3 | –2.107 |
| | **Ratio ~ taxonomic group + ratio type + (1 \| StudyID / PopulationID)** | 13 | –102685.1 | –0.881 |
| Fifty | Fifty ~ taxonomic group + marker + loci number + (1 \| StudyID) | 12 | 3607.0 | 0.000 |
| | Fifty ~ taxonomic group + marker + (1 \| StudyID) | 11 | 3606.5 | –0.518 |
| | **Fifty ~ taxonomic group + (1 \| StudyID)** | 10 | 3607.7 | 0.659 |
| Fivehundred | Fivehundred ~ taxonomic group + marker + loci number + (1 \| StudyID) | 12 | 2209.4 | 0.000 |
| | **Fivehundred ~ taxonomic group + marker + (1 \| StudyID)** | 11 | 2207.4 | –2.011 |
| HFI | $N_e$ ~ taxonomic group * HFI + marker + loci number + (1 \| StudyID / PopulationID) | 24 | 2004481 | 0.000 |
| | **$N_e$ ~ taxonomic group * HFI + marker + (1 \| StudyID / PopulationID)** | 23 | 2004479 | –2.024 |

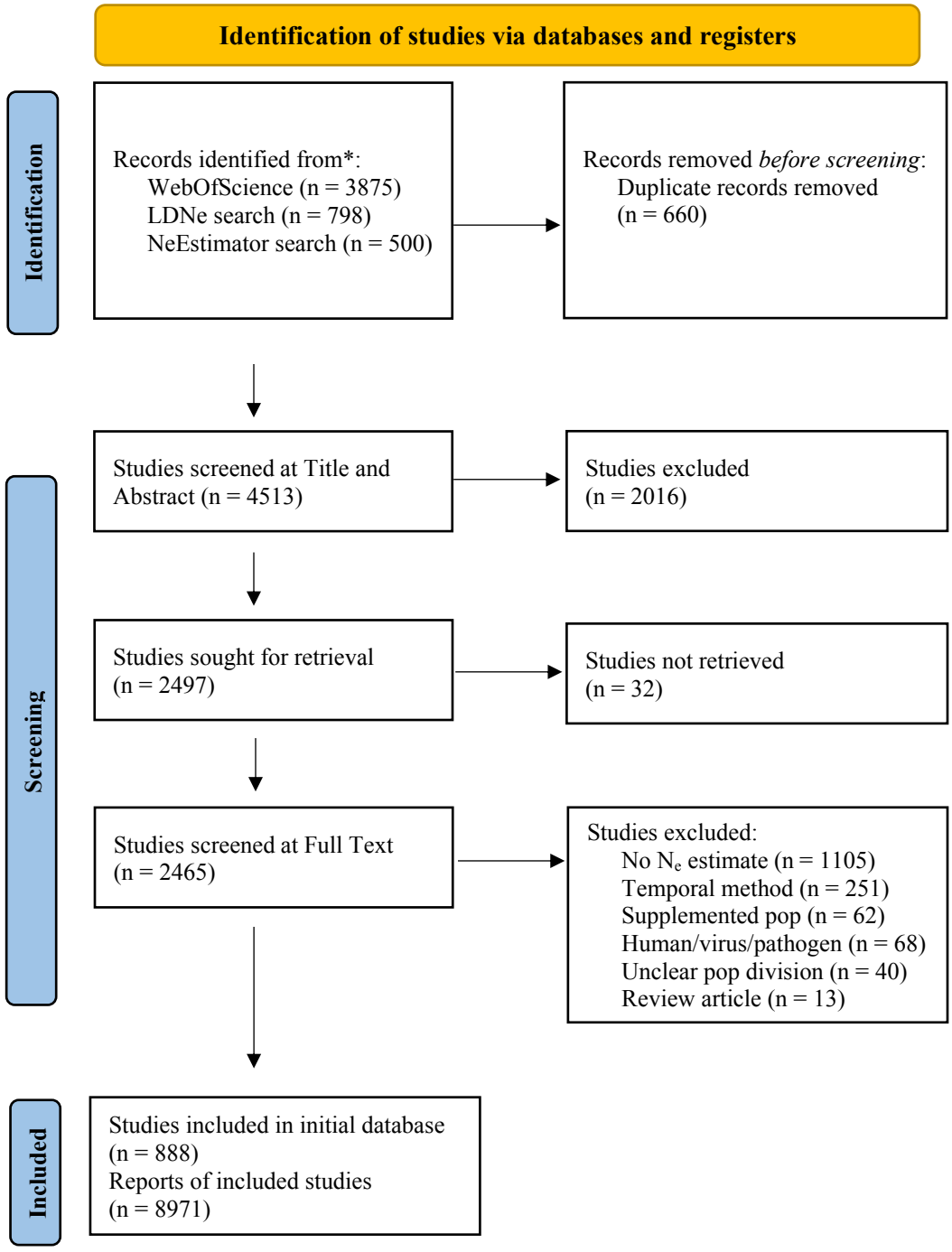


**Figure S1**: Flow diagram of the review process, modified from O'Dea et al. (2021)

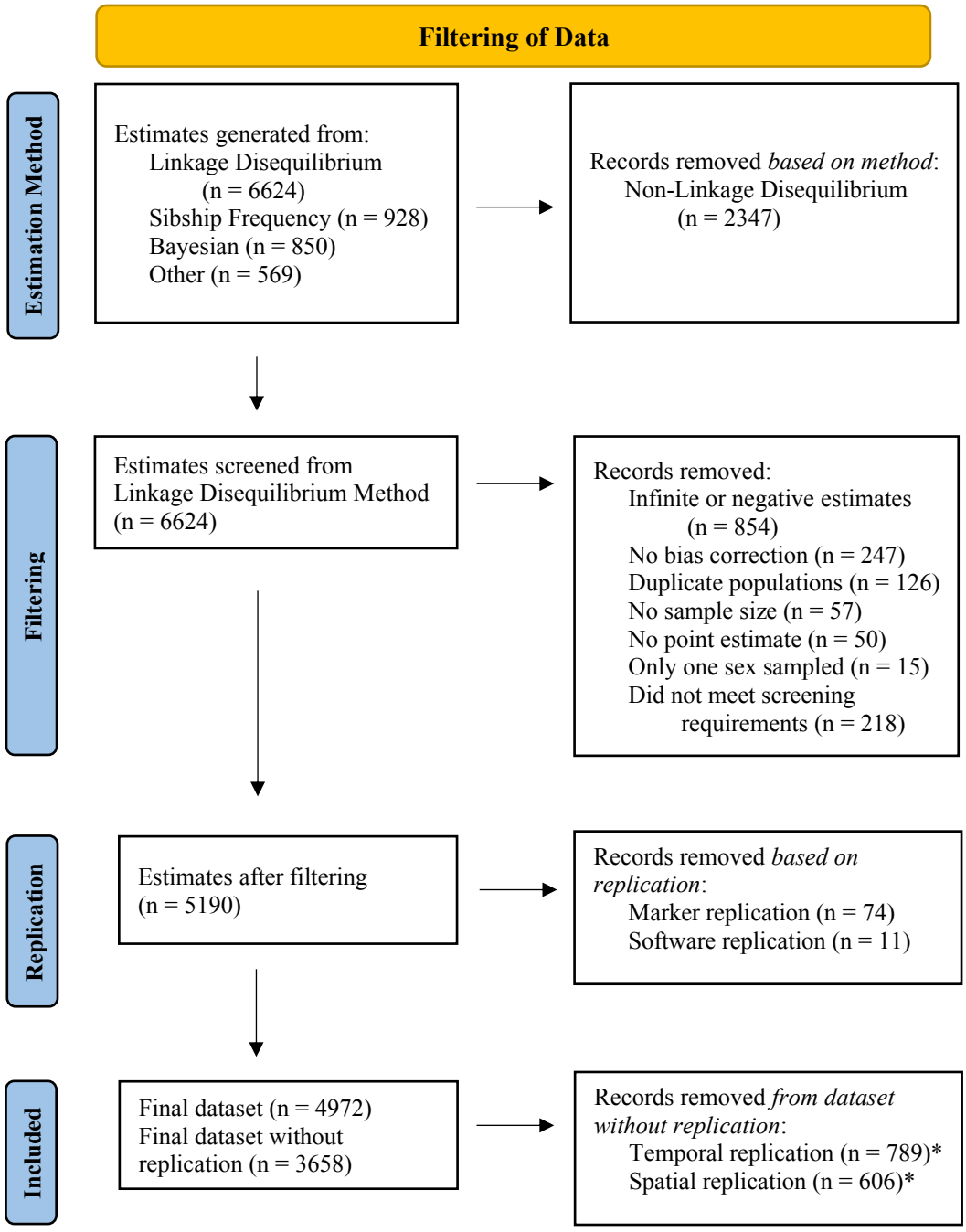


**Figure S2**: Flow diagram of the filtering process, modified from O'Dea et al. (2021). *Some records had both spatial and temporal replication

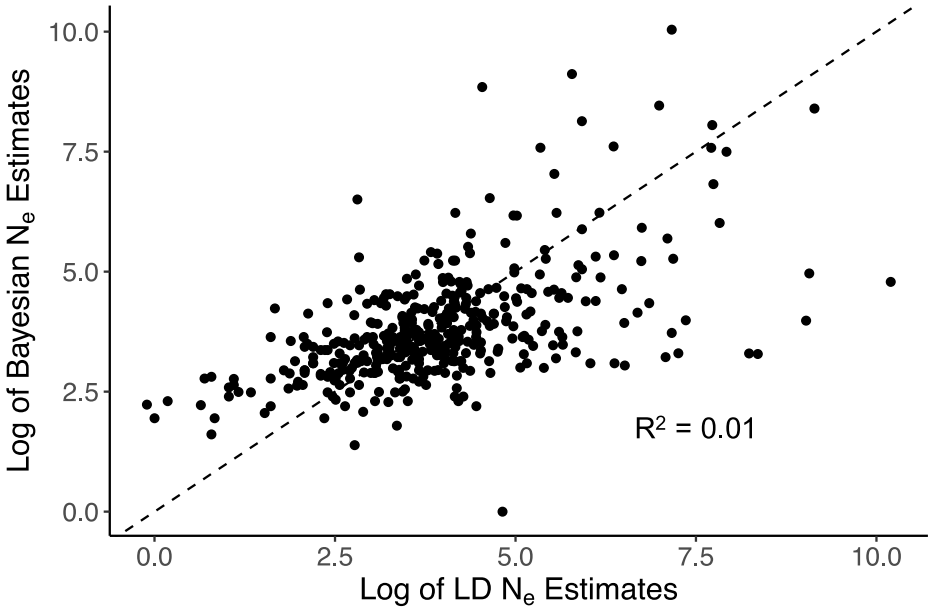


**Figure S3**: Regression between $N_e$ estimates produced from the Linkage Disequilibrium (LD) method, and the Bayesian Computation method. Sample size for the analysis was N = 416.

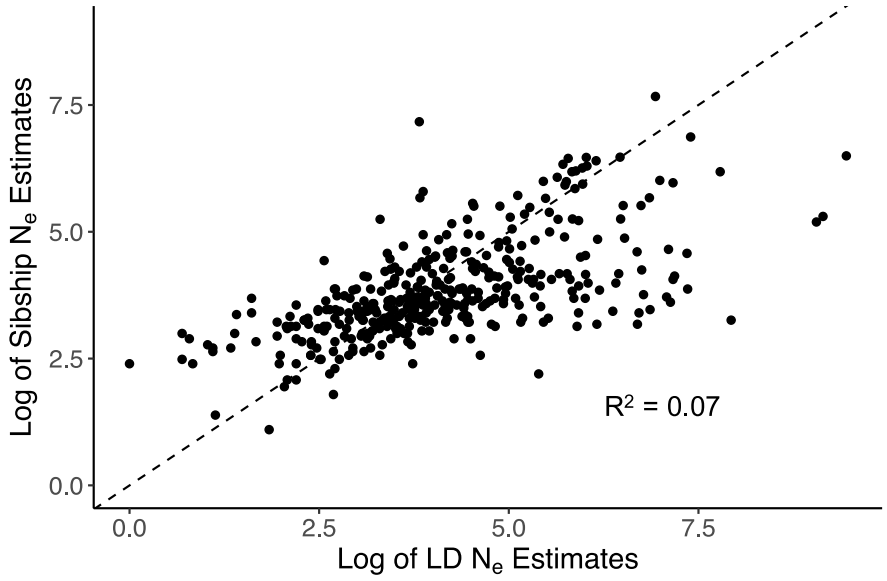


**Figure S4**: Regression between $N_e$ estimates produced from the Linkage Disequilibrium (LD) method, and the Sibship Frequency method. Sample size for the analysis was N = 386.

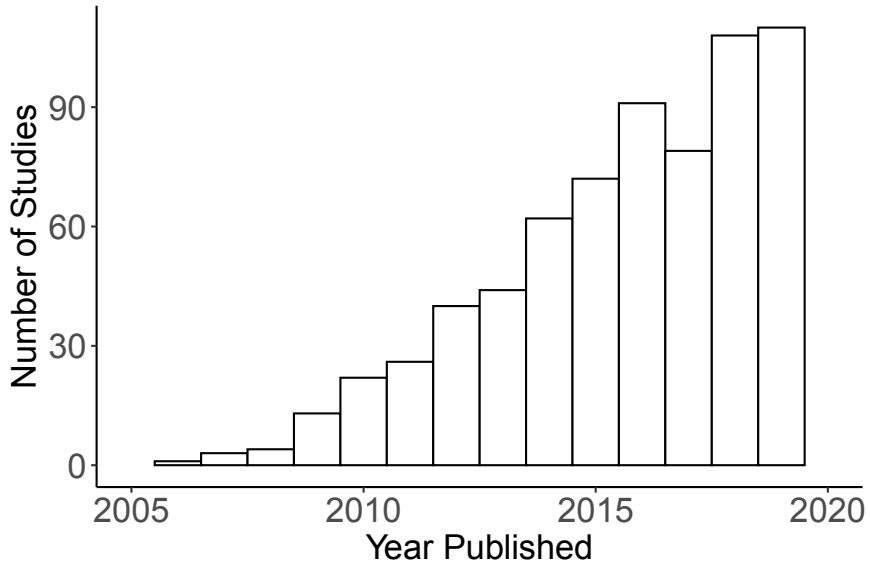


**Figure S5**: Number of studies with $N_e$ or $N_b$ estimates published per year. Studies from 2020 were not included as the literature search was conducted partway through the year and is not representative of all articles published

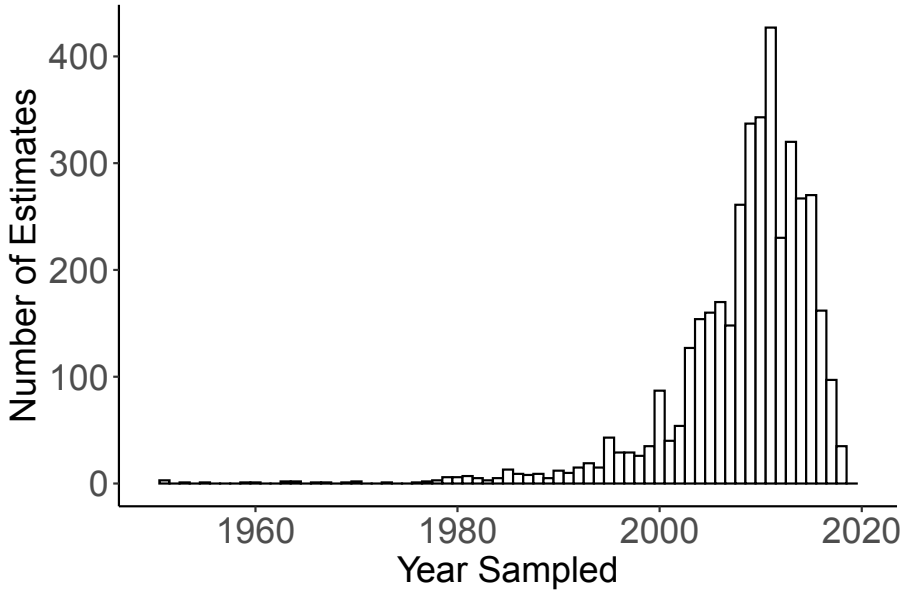


**Figure S6**: Number of estimates of $N_e$ or $N_b$ sampled per year in the final dataset. There were 957 estimates with no sampling year reported.

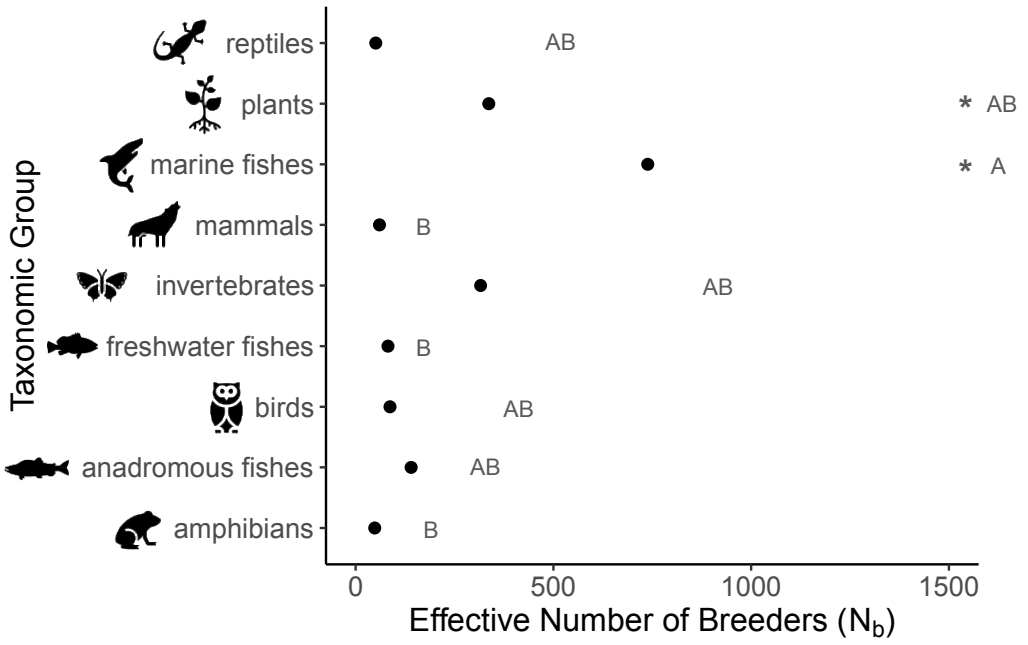


**Figure S7**: Mean effective number of breeders ($N_b$) across taxonomic groups, accounting for differences in marker type. Shaded green bars represent 95% confidence intervals. *The upper CI for plants and marine fish are not contained within the plot. Groups with no shared letters are statistically different from one another.

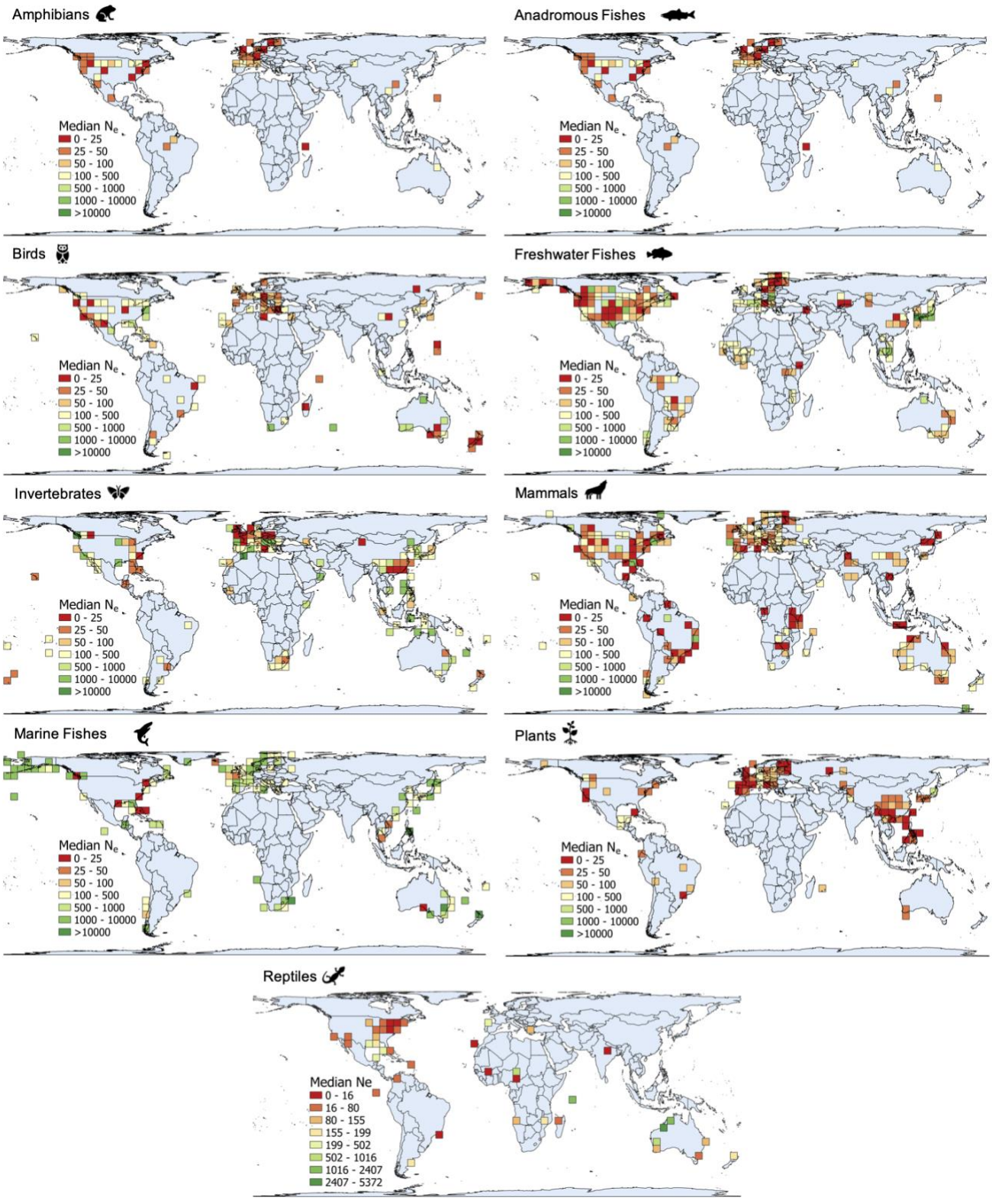


Figure S8: Global maps showing the taxa-specific median $N_e$ value in each 500×500km grid cell. Data is projected with the world Behrmann projection