University of Massachusetts Amherst

Reproduction in the Wild: The Effect of Individual Life History Strategies on Population Dynamics and Persistence

Item Type	dissertation
Authors	Coombs, Jason Asa
DOI	10.7275/1667664
Download date	2025-01-06 02:15:05
Link to Item	https://hdl.handle.net/20.500.14394/38708

REPRODUCTION IN THE WILD: THE EFFECT OF INDIVIDUAL LIFE HISTORY STRATEGIES ON POPULATION DYNAMICS AND PERSISTENCE

A Dissertation Presented

by

JASON ASA COOMBS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY
September 2010
Graduate Program in Organismic & Evolutionary Biology

© Copyright by Jason Asa Coombs 2010
All Rights Reserved

REPRODUCTION IN THE WILD: THE EFFECT OF INDIVIDUAL LIFE HISTORY STRATEGIES ON POPULATION DYNAMICS AND PERSISTENCE

A Dissertation Presented

by

JASON ASA COOMBS

Approved as to style and content by:

_____
Benjamin H. Letcher, Chair

_____
Keith H. Nislow, Member

_____
Jeffrey Podos, Member

_____
Andrew R. Whiteley, Member

_____
Elizabeth Dumont, Director
Graduate Program in Organismic & Evolutionary Biology

DEDICATION

To my wife and sons, for providing perspective

ACKNOWLEDGMENTS

I would like to thank Ben Letcher and Keith Nislow for their guidance, insight, patience, and (in my opinion) superb ability to secure funding to make this research, and this dissertation, possible. I would also like to thank Kevin McGarigal, Jeff Podos, and Andrew Whiteley for providing advice and insight on this research.

I would like to thank Matt O'Donnell and Todd Dubreuil for their assistance with all things West Brook, Mark Hudy and Gonzalo Mendez for all things Fridley Gap, and Tim King for sharing his brook trout primers. I would like to thank the following people who helped with fieldwork (In alphabetical order): Carol Coombs, Tyler Evans, Gregg Horton, Mariska Obedzinski, Jeff Ojala, Jamie Pearlstein, Justin Scace, Doug Sigourney, Jim Sotiropoulis, Amy Varady, and a myriad of volunteers.

I would like to thank all of those involved with the OEB program, particularly everyone who volunteered to conduct annual student reviews, and especially Penny Jacques for being a terrific facilitator throughout the entire process.

No research is possible without funding, and thus I would like to thank the following agencies for monetary support: USFS Northern Research Station, USGS Conte Anadromous Fish Research Center, USFS Fish & Aquatic Ecology Unit, and The Nature Conservancy.

Lastly, and mostly, I would like to thank the brook trout for allowing me a glimpse into their complex and fascinating life-histories. To any that I inadvertently, or advertently, killed, I am truly sorry.

ABSTRACT

REPRODUCTION IN THE WILD: THE EFFECT OF INDIVIDUAL LIFE HISTORY STRATEGIES ON POPULATION DYNAMICS AND PERSISTENCE

SEPTEMBER 2010

JASON ASA COOMBS, B.A., LOCK HAVEN UNIVERSITY

M.A., CLARION UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Benjamin H. Letcher

For a sexually reproducing species, the two major decisions facing all individuals are when and with whom to reproduce. When scaled to the population level, the outcome from all individual decisions determines reproductive variance, and age-class contribution to population growth rate. Both of these attributes determine a population's effective size (N_e), which is directly correlated with its fitness, persistence probability, and adaptability.

The questions of when and with whom to reproduce, and their subsequent effects on N_e and age-at-maturity were assessed for wild brook trout (*Salvelinus fontinalis*) populations. Mating pairs were significantly size-assortative, with individual length accounting for 37% of the variation. This pattern of size assortative mate choice resulted in a reproductive strategy closer to monogamy than polygamy. Of all reproducing adults (n=157), 80% (n=126) produced only one full-sibling family, and only 6% (n= 9) contributed to more than two full-sibling families. The number of families and offspring contributed increased with length for both males and females. Comparison of the

effective population size estimate to the adult census size (N_c) estimate returned an N_e:N_c ratio of 0.49 averaged over both populations. This value is nearly five times greater than the average reported across 165 (0.14) and 102 (0.10) different species.

Age-at-maturity ranged from 0 to 2 years, with the proportion of age-0 and age-1 individuals maturing in a given year dependent upon growth opportunities determined primarily by environmental conditions. Mature fish were significantly larger than immature fish within an age-class, however, survival rates of mature and immature fish were similar. Furthermore, parental length did not influence offspring survival. These data suggest that the cost of early maturation is instead manifested through a reduction in egg number for females, and a reduced ability to acquire mates for males, both determined by an individual's size. Indeed, fecundity predicted by mean length of immature and mature fish within an age-class would result in mature fish producing an average of 38% (age-0) and 33% (age-1) more eggs than immature fish.

These findings are discussed in the context of population persistence given the trend of increasing habitat fragmentation and looming climate change.

TABLE OF CONTENTS

Page

AKNOWLEDGEMENTS……. v

ABSTRACT…….. vi

LIST OF TABLES…… xi

LIST OF FIGURES……... xii

I. PEDIGREE SIMULATION AND RECONSTRUCTION SOFTWARE……... 1

CREATE: Software to create input files from diploid genotypic data for 52 genetic software programs……...…….. 1

Abstract…….…… 1
Program Description…….. 1

PEDAGOG: Software for simulating eco-evolutionary population dynamics…… 5

Abstract…….. 5
Program Description…….. 5

PedAgree: Software to quantify error and assess Accuracy and congruence for genetically reconstructed pedigree relationships..... 16

Abstract…….. 16
Program Description…….. 16

II. GENETICALLY RECONSTRUCTED PEDIGREES: THE COSTS AND BENEFITS OF USING FULL-SIBLING STRUCTURE TO CONSTRAIN PARENTAGE ASSIGNMENTS…….…24

Abstract…….. 24
Introduction……....25
Materials and Methods……...28

Simulations…… 28
Pedigree Reconstruction…… 31
Analyses……33

Results……34

Analysis Variables…… 36
Population Variables……..36
Power Variables……37
Error Variables……..……38

Discussion……..38

III. MATING PATTERNS IN WILD BROOK TROUT POPULATIONS: INFLUENCES ON EFFECTIVE POPULATION SIZE AND IMPLICATIONS FOR POPULATION PERSISTENCE……... 71

Abstract……..71
Introduction……72
Materials and Methods……...75

Study Sites and Sampling Design……..… 75
Genetic Diversity and Genotyping Error……... 77
Pedigree Reconstruction…… 79
Census and Effective Population Size Estimation……80

Results……82

Genetic Diversity and Genotyping Error……... 82
Pedigree Reconstruction…… 83
Census and Effective Population Size Estimation……85

Discussion……..87

IV. WHEN TO REPRODUCE: CAUSES AND CONSEQUENCES OF EARLY MATURATION IN AN INDETERMINATELY GROWING SPECIES……. 109

Abstract……..109
Introduction…….....110
Materials and Methods……...113

Study Site and Sampling Design……... 113
Genetic Diversity and Genotyping Error……... 114
Pedigree Reconstruction…… 115
Environmental Variables……... 116
Population Densities and Length Metrics…….....117
Survival Analysis…….....118

Results……119

Genetic Diversity and Genotyping Error……... 119
Pedigree Reconstruction…… 120
Environmental Variables……... 121
Population Densities and Length Metrics…….. 121
Survival Analysis……... 124
Age-Class Contribution……. 125

Discussion…….. 126

V. SYNOPSIS……..... 149

APPENDIX: ALLELE FREQUENCIES FOR CHAPTER II SIMULATIONS…….. 152

BIBLIOGRAPHY……155

LIST OF TABLES

Table Page

2.1. Diversity measures of microsatellite loci used in simulations……50

2.2. Parentage assignment classifications and their use in accuracy assessment measures……..51

2.3. Average mean and 95% confidence interval for percent differences in parentage assignment accuracies between the sibship constraint method and parentage assignment algorithms…… 52

2.4. Effect of genetic sibship reconstruction program and minimum threshold value on sibship constrained and raw parentage assignment accuracies……... 53

2.5. Effect of mating strategy and mean sampled full-sibling family size on sibship constrained and raw parentage assignment accuracies……..56

2.6. Effect of proportion of actual parents sampled and number of loci used in analyses on sibship constrained and raw parentage assignment accuracies…….. 59

2.7. Effect of genetic and demographic error rates on sibship constrained and raw parentage assignment accuracies……..65

3.1. Single locus summary statistics for West Brook and Fridley Gap brook trout cohorts……... 93

3.2. Pedigree descriptive measures for reconstructed full-sibling families containing a minimum of two individuals for brook trout populations inhabiting the West Brook and Fridley Gap drainages…….. 99

3.3. Estimates of effective population size generated by single-sample and temporal methods for brook trout populations inhabiting the West Brook and Fridley Gap drainages……. 100

4.1. Single locus summary statistics for brook trout cohorts inhabiting the open-small and open-large populations……...….. 133

4.2. Pedigree descriptive measures for reconstructed full-sibling families containing a minimum of two individuals for brook trout populations inhabiting the open-small and open-large tributaries……137

4.3. Average daily discharge and temperature over the spring and summer seasons, and number and density of brook trout captured during the fall sample in the open-small (OS) and open-large (OL) tributaries……138

LIST OF FIGURES

Figure Page

1.B.1. Linear fit for offspring trait value regressed against mid-parental trait value for the tenth generation one year after birth…… 14

1.B.2. The affects of heritability and selection on trait value distribution for a single cohort after ten generations…… 15

1.C.1. An example of a sibship reconstruction accuracy summary file…….22

1.C.2. An example of a parentage assignment accuracy summary file……. 23

3.1. Map of study regions…….. 103

3.2. Histograms of full-sibling family sizes by cohort for brook trout populations inhabiting the West Brook and Fridley Gap drainages…….. 104

3.3. Individual contribution of full-sibling families by parents for brook trout inhabiting the West Brook and Fridley Gap drainages……..105

3.4. Individual full-sibling family contributions as a function of individual length for brook trout populations inhabiting the West Brook and Fridley Gap drainages……106

3.5. Total offspring contributed by an individual as a function of individual length for brook trout populations inhabiting the West Brook and Fridley Gap drainages……107

3.6. Size of mating pairs in the West Brook and Fridley Gap brook trout populations…….....108

4.1. Map showing the West Brook drainage located in Whately, MA, USA…… 139

4.2. Sizes of young-of-year brook trout caught in the fall sample of 2002, 2003, and 2004 in the open-small and open-large tributaries…… 140

4.3. Sizes of age-1 brook trout caught in the fall sample of 2002, 2003, and 2004 in the open-small and open-large tributaries……..…… 141

4.4. Mean size of immature and mature age-0 and age-1 brook trout inhabiting the open-small and open-large tributaries during the 2002-2004 spawning years..... 142

4.5. Length in spring, and spring-to-fall growth rate of immature and mature age-1 brook trout during the 2003 and 2004 spawning years……143

4.6. Condition factor in the spring and summer for age-1 immature and mature brook trout during the 2003 and 2004 spawning years……..144

4.7. Annual survival estimates for immature and mature brook trout detected during the 2002-2004 spawning years…….…….. 145

4.8. Effect of maternal and paternal lengths on offspring survival……146

4.9. Ages of contributing parents for the 2002-2004 brook trout cohorts inhabiting the open-small and open-large tributaries……147

4.10. Proportion of offspring contributed by parental age-class for the 2002-2004 spawning years in the open-small and open-large brook trout populations…..148

CHAPTER I

PEDIGREE SIMULATION AND RECONSTRUCTION SOFTWARE

CREATE: Software to create input files from diploid genotypic data for 52 genetic software programs

Abstract

CREATE is a Windows program for the creation of new and conversion of existing data input files for 52 genetic data analysis software programs. Programs are grouped into areas of sibship reconstruction, parentage assignment, genetic data analysis, and specialized applications. CREATE is able to read in data from text, Microsoft Excel and Access sources and allows the user to specify columns containing individual and population identifiers, birth and death data, sex data, relationship information, and spatial location data. CREATE's only constraints on source data are that one individual is contained in one row, and the genotypic data is contiguous. CREATE is available for download at https://bcrc.bio.umass.edu/pedigreesoftware/.

Program Description

The proliferation of analyses involving usage of codominant, diploid genotypic markers such as microsatellites has led to the availability of a myriad of software

programs. Unfortunately, many programs require specific input file formats that must be created using a text editor and are not readily reproduced if data either changes or is updated. Depending on the intricacies of the input file format specifications, creation of these files can consume vast amounts of time, introduce errors into the data, and may even deter the use of a program altogether. To allay these concerns, some programs have built in the ability to convert between input file formats (for a detailed schematic see Figure 1 in Excoffier & Heckel (2006)), and two programs were designed specifically for the purpose of creation and conversion of input files (Glaubitz 2004; Manoukis 2007). Unfortunately these programs almost exclusively convert input file formats only to multi-purpose genetic data analysis programs, leaving out programs designed for parentage analysis, sibship reconstruction, and many specialized applications. Additionally, it is still oftentimes necessary to use several programs to reach the desired input file format. For example, to convert a file formatted for ARLEQUIN into a file formatted for FSTAT, it would be necessary to use an intermediate file in GENEPOP format. In addition to taking time, this increases the potential to incorporate errors and lose information such as population names and individual identifiers.

We have developed a program that creates input files for 52 software programs from raw data. The programs are grouped into categories of sibship reconstruction, parentage analysis, multi-purpose genetic data analysis, and specialized applications. Two major advantages of CREATE are 1) raw data for input file creation can be accessed from text files delimited with any character, any spreadsheet in a Microsoft Excel workbook, or any table or query in a Microsoft Access database, and 2) the only constraints on the raw data format are that all data for an individual must be contained in

one row, and the genotypic data must be contiguous, with alleles for a loci located in adjacent columns. We have strived to make the program as flexible as possible by allowing the user the option of including additional information such as population and individual identifiers, birth and death data, sex data, known parent-offspring relationships, and spatial location information located in columns on either side of the genotypic data. Additionally, the raw data may begin in any row, may include a row containing column titles, and is able to use any character to designate missing allele values.

Conversion from preexisting input files is also possible for certain programs, provided that enough information is present. For example, conversion from GENEPOP to PEDIGREE would be allowed, but conversion from GENEPOP to PASOS would not be possible due to a lack of information to differentiate parents from offspring. We strongly recommend creating input files from raw data whenever possible to decrease information loss and error propagation.

To date, input files can be created from raw data for the following programs: ARLEQUIN (Excoffier *et al*. 2005), BAPS4 (Corander & Marttinen 2006), BATWING (Wilson *et al*. 2003), BAYESASS (Wilson & Rannala 2003), BOTTLENECK (Piry *et al*. 1999), CERVUS (Marshall *et al.* 1998), COLONISE (Foll & Gaggiotti 2005), COLONY (Wang 2004), CONE (Anderson 2005), FAMOZ (Gerber *et al*. 2003), FAMSPHERE (Carvajal-Rodriguez 2007), FDIST2 (Beaumont & Nichols 1996), FSTAT (Goudet 1995), GDA (Lewis & Zaykin 2001), GENECLASS2 (Piry *et al*. 2004), GENEPOP (Raymond & Rousset 1995), GENETIX (Belkhir *et al*. 2004), GERUD2 (Jones 2001; Jones 2005), GIMLET (Valiere 2002), IDENTIX (Belkhir *et al*. 2002), IM/IMa (Hey &

Nielsen 2004; Hey & Nielsen 2007), IMMANC (Rannala & Mountain 1997), KINGROUP (Konovalov *et al*. 2004), LAMARC (Kuhner 2006), MICRO-CHECKER (Van Oosterhout *et al*. 2004), MICROSAT (Minch *et al.* 1996), MIGRATE (Beerli 2006), MLNE (Wang & Whitlock 2003), ML-RELATE (Kalinowski *et al*. 2006), MSA (Dieringer & Schlotterer 2003), MSVAR (Beaumont 1999), NEESTIMATOR (Peel *et al*. 2004), NEWHYBRIDS (Anderson & Thompson 2002), NEWPAT (Wilmer *et al.* 1999), PAPA (Duchesne *et al*. 2002), PARENTAGE (Emery *et al.* 2001), PARENTE (Cercueil *et al*. 2002), PASOS (Duchesne *et al*. 2005), PEDAPP (Almudevar 2007), PEDIGREE (Butler *et al*. 2004;Smith *et al*. 2001), PHYLIP (Felsenstein 2004), PROBMAX3 (Danzmann 1997), PRT (Almudevar & Field 1999), SALMONNB (Waples *et al*. 2007), SPAGEDI (Hardy & Vekemans 2002), STRUCTURE (Falush *et al*. 2003), TFPGA (Miller 1997), TM3 (Berthier *et al*. 2002), TMVP (Beaumont 2003), WHICHLOCI (Banks *et al*. 2003), WHICHPARENTS (Hedgecock & Eichert 1999) and WHICHRUN (Banks & Eichert 2000).

An executable version of CREATE along with documentation and example data files can be downloaded at https://bcrc.bio.umass.edu/pedigreesoftware/. We intend to remain active in implementing new software program options as future programs are created or as a need arises for existing ones.

PEDAGOG: Software for simulating eco-evolutionary population dynamics

Abstract

PEDAGOG is a Windows program that can be used to determine power for, and validate inferences drawn from, eco-evolutionary studies. It models dynamics of multiple populations and their interactions through individual based simulations while simultaneously recording genotype, pedigree, and trait information at the individual level. PEDAGOG also allows for specification of heritable traits, natural and sexual selection acting upon those traits, population sampling schemes, and incorporation of genetic and demographic errors into the output. Overall, parameters can be specified for genetic diversity, demographics, mating design, genetic and demographic errors, individual growth models, trait heritability and selection, and output formatting. Demographic parameters can be either age or function based, and all parameters can be drawn from twelve statistical distributions where appropriate. Simulation results can be automatically formatted for 57 existing software programs to facilitate post-simulation analyses. PEDAGOG is freely available for download at https://bcrc.bio.umass.edu/pedigreesoftware/.

Program Description

Increasingly, studies of natural populations integrate aspects of genetics, ecology, and evolution to investigate eco-evolutionary processes and dynamics (Hairston *et al.*

2005; Hanski & Saccheri 2006; Kinnison & Hairston 2007; Kruuk & Hill 2008; Saccheri & Hanski 2006). Because of the complex nature of these processes, the high level of variation occurring in most natural systems and populations, and the fact that datasets are typically incomplete and contain errors, it becomes imperative to assess inferences drawn from the empirical data (Morrissey *et al.* 2007; Pemberton 2008).

One approach for inference evaluation would involve the use of synthetic data. The process would involve the simulation of datasets that possess the characteristics of the empirical system under investigation. Furthermore, the simulation model should be parameterized with the empirically derived values of interest. The synthetic data would then be subjected to the same methodology used on the empirical data to determine if the model parameter values were recovered. This would enable investigators to assess the reliability and robustness of inferences.

To generate such synthetic data for inference evaluation would often necessitate the use of a simulation program that can perform individual based modeling and record individual genotype, pedigree, and trait information while allowing for the incorporation of eco-evolutionary processes. Although development of one's own simulation program is an option, the complexity of eco-evolutionary processes and their interactions would result in a significant time commitment and high potential for error. Another option would be to use an existing program that has already undergone validation.

Unfortunately, the majority of programs developed for eco-evolutionary processes have focused on data analysis rather than population simulation (Coombs *et al.* 2008; Excoffier & Heckel 2006). Of the simulation programs available, only EASYPOP (Balloux 2001) and SPIP (Anderson & Dunham 2005) record individual pedigree and

genotype information. However, EASYPOP does not allow for specification of a population sampling scheme, SPIP can only simulate a single population, and neither allow for the addition of heritable traits that can be used for both natural and sexual selection. Therefore we developed PEDAGOG, a user-friendly, flexible program for the realistic simulation of population dynamics for multiple populations that allows for evolutionary processes while recording individual pedigree, genotype, and trait information.

PEDAGOG can be used to assess conclusions drawn from empirical data, or determine the feasibility of potential studies for numerous questions regarding topics such as pedigree reconstruction accuracy, trait heritability, natural selection, sexual selection, effective population size, capture-mark-recapture, population structuring, population viability analysis, life-history strategy, inbreeding depression, dispersal, density-dependence, and optimality modeling to name a few. For example, simulated data could be used to assess the effect of incorrect parentage assignments on the estimate of trait heritability under differing selection strengths for a given genotypic marker set. Figure 1.B.1 shows that the magnitude of the effect of pedigree error decreases as the strength of selection increases. Another use of PEDAGOG could be to predict trait distributions under varying levels of heritability and selection. A scenario involving three levels of negative directional selection shows that heritability affects the variance of the trait's distribution while selection influences the mean (Figure 1.B.2).

PEDAGOG can simulate up to fifteen populations and their interactions concurrently. Within each population, parameters are divided into seven primary areas each represented by a tab located on the main form. All parameters are initially filled

with default values enabling the user to modify only the areas of interest for the current simulation. Parameters can either be specified by hand or recalled from a previously saved file. Specification for the majority of parameters allows them to be drawn from one of twelve statistical distributions, allowing for great flexibility in recreating the desired distribution. Saving parameters to or recalling parameters from a file can be performed for all tabs concurrently or for each tab singly. There is a tab for each of the following seven areas: genetics, demographics, mating design, mutation/error, individual growth, heritability/selection, and output.

The ’Genetics’ tab contains parameters specifying genetic marker information. PEDAGOG deals explicitly with diploid markers, and was designed specifically for microsatellites. Up to 48 loci may be specified, each with up to 90 alleles. Loci number and allele frequencies may be specified by hand or imported from raw genotypic data. Additional information specified for each locus includes name, repeat length, which populations it’s scored for, pair-wise linkage with other loci, and allele specific null-allele presence.

The ‘Demographics’ tab incorporates population parameters for size, location, and pair-wise emigration and immigration probabilities. Population sizes can be constant, random, density dependent, geometric, or drawn from a distribution. There are parameters for bottleneck occurrences and sex proportions, as well as individual probabilities for maturation, movement, capture probability, and survival, all of which can be either age or function based. There are options to incorporate density dependence effects into survival, growth, and movement probabilities. Additional parameters specified include number of samples between reproductive events and their time of

occurrence, number of generations to simulate, and whether to begin simulations from population pools or continue from a previous PEGAGOG simulation.

The ‘Mating Design’ tab contains parameters for mating strategy, mate choice, mate number, and fecundity. Mating strategy can be monogamous, polygynous, polyandrous, or polygamous. There is an option to incorporate sexual selection into mate choice, along with the range of cohorts able to reproductively overlap. There are variables for fecundity and mate number which can be either age or function based. If the number of potential sires is greater than one, sire contribution can be specified as uniform, random, size or age proportional, or size or age dominant.

The ‘Mutation/Error’ tab contains parameters for genetic mutations and genetic and demographic errors. Genetic mutation parameters specify probabilities of primer site (null allele) and allelic mutation occurrence for each locus (Ewen *et al.* 2000; O'Reilly *et al.* 1998). Allelic mutations can follow an infinite allele model (Kimura & Crow 1964), stepwise mutation model (Ohta & Kimura 1973), or a combination of both. Genotyping error rates are also locus specific and can be specified for large allele dropout (Wattier *et al.* 1998) and allele miscall. A miscall can be classified as an adjacent allele, a false allele, or a combination of both.

Demographic errors incorporate incorrect information into the output, and are meant to replicate the types of errors that arise in a typical dataset. Probability of erroneous cohort (age) assignment can be specified based on an individual’s age or size. Both scenarios allow for an individual to be misclassified into an adjacent cohort based on its size relative to the cohort mean. Additional demographic error parameters include sex misclassification and incorrect parent identification.

The ‘Individual Growth’ tab offers the choice of six individual growth models (Gamito 1998): exponential, restricted, logistic, parabolic, Gompertz, and von Bertalanffy. The attribute modeled can be either length or weight, and a relationship can be established between the two. Growth model parameters can be specified for groups ranging from all individuals to ones classified by all combinations of sex and maturity status. Additionally, there is a seasonal growth function that specifies the timing of growth accumulation between reproduction events.

The ‘Heritability/Selection’ tab allows for specification of heritability and selection types and strengths for individual size, age at maturation, and movement, along with up to ten custom traits. The source of the trait’s heritability may be maternal, paternal, or parental, and its strength can range from zero for no heritability to one for complete heritability. The type of selective pressure may be directional, stabilizing, or disruptive, and strengths can range from zero to one (negative one for negative directional selection). The strength alters individual survival probability based upon the individual’s trait value and the accompanying selection function. Selection types and strengths can be specified for multiple traits, with the final affect on survival probability averaged over all traits.

The ‘Output’ tab contains parameters determining the final content of the output, along with program selections determining which input files will be created from the content for subsequent analysis. Output parameters specify the proportion of individuals having known sex, known parents, and known timing of death. Additional parameters specify the probability an individual will be genotyped for a specific locus, and whether

captured individuals are genotyped on every capture event, or only their initial one. This area also specifies which generations to sample, and the number of replications to run.

The output file area also provides users with the option of selecting post-simulation analysis programs for which to automatically create input files for from the simulated data. Fifty-two of the programs are listed in Coombs *et al.* (2008) and described in the program CREATE's user guide. An additional five programs have been added for areas of quantitative genetics (WOMBAT (Meyer 2006), VCE (Neumaier & Groeneveld 1998)), capture-mark-recapture (MARK (White & Burnham 1999), M-SURGE (Choquet *et al.* 2004)), and isolation by distance (IBD (Bohonak 2002). There is also an option to create a batch file that is used by the program PEDAGREE (Coombs *et al.* 2010b) to simplify and speed-up the process of sibship reconstruction and parentage assignment program accuracy assessment.

In addition to input files formatted for external programs, PEDAGOG produces five other output files for each simulation repetition. (1) A 'Complete Pedigree' file containing all data for all individuals. This file is also used to initialize a new simulation from a previous simulation allowing the previous run to be extended. (2) A 'Null Alleles' file that is used in conjunction with the 'Complete Pedigree' file for continuation of a simulation. (3) A 'Complete Sampling' file that records data for all captured individuals and is used for capture-mark-recapture analysis. (4) A 'True Genotypes' file that records actual trait and genotype data for captured individuals. And (5) an 'Apparent Genotypes' file which is the genetic and demographic error containing version of the 'True Genotypes' file. All input files are created using data in the 'Apparent Genotypes' file with the exception of those involved with capture-mark-recapture analysis which are

created using data from the 'Complete Sampling' file. Data in the 'True Genotypes' file is used for comparative purposes in error evaluation.

PEDAGOG was validated by two different methods. The first compared the loss of observed heterozygosity values over the period of one hundred generations to predicted heterozygosity values computed from mean effective population size and inbreeding coefficient measures using the following formula (Falconer & Mackay 1996):

$$\frac{H_t}{H_0} = \left(1 - \left(\frac{1}{2N_e}\right)\right)^t = 1 - F$$

where H_0 is the initial heterozygosity, H_t is the observed heterozygosity at generation t, N_e is the effective population size, t is the number of generations, and F is the mean inbreeding coefficient of the population. Effective population sizes were estimated using the programs MLNe (Wang 2001; Wang & Whitlock 2003), and LDNe (Waples & Do 2008). The average inbreeding coefficient was calculated using the program PEDIG (Boichard 2002). Results showed observed heterozygosity of the simulated population in almost identical agreement with predicted values based on both effective population size estimates and inbreeding coefficient calculations verifying that genotypic data were simulated correctly.

The second validation method tested population interactions by comparing observed and predicted allele frequencies for an island model scenario. Predicted allele frequencies were calculated using equation 21 from Nagylaki (1979)

$$\overline{\mathrm{X}}_t = \overline{\xi} + \left(\chi_0 - \overline{\xi}\right)\left(1 - \overline{m}\right)^t$$

where $\overline{\mathrm{X}}_t$ is the predicted allele frequency of sub-population X, t is the time in generations, $\overline{\xi}$ is the mean allele frequency of all sub-populations, χ_0 is the initial allele

frequency of sub-population X, and $\overline{m}$ is the mean migration rate of all sub-populations. Simulations were conducted using five sub-populations with initial frequencies of allele *A* at a biallelic locus set to 1 (Pop 1), 0.75 (Pop 2), 0.5 (Pop 3), 0.25 (Pop 4), and 0 (Pop 5), $\overline{m}$ equal to 0.1, and *t* set to 10. Predicted allele frequency values for all five sub-populations were contained within the 95% confidence interval estimates of mean observed allele frequencies calculated from fifty replications. This verifies that population interactions are simulated correctly.

We intend to remain active in continued development of PEDAGOG and specifically would like to add options allowing the linking of genetic markers to traits thus enabling QTL analysis, incorporate individual inbreeding coefficient's to enable selection against inbred individuals, add in a dynamic energy budget growth model, and expand the number and type of genetic markers available. An executable version of PEDAGOG along with documentation and example project files can be freely downloaded at https://bcrc.bio.umass.edu/pedigreesoftware/.

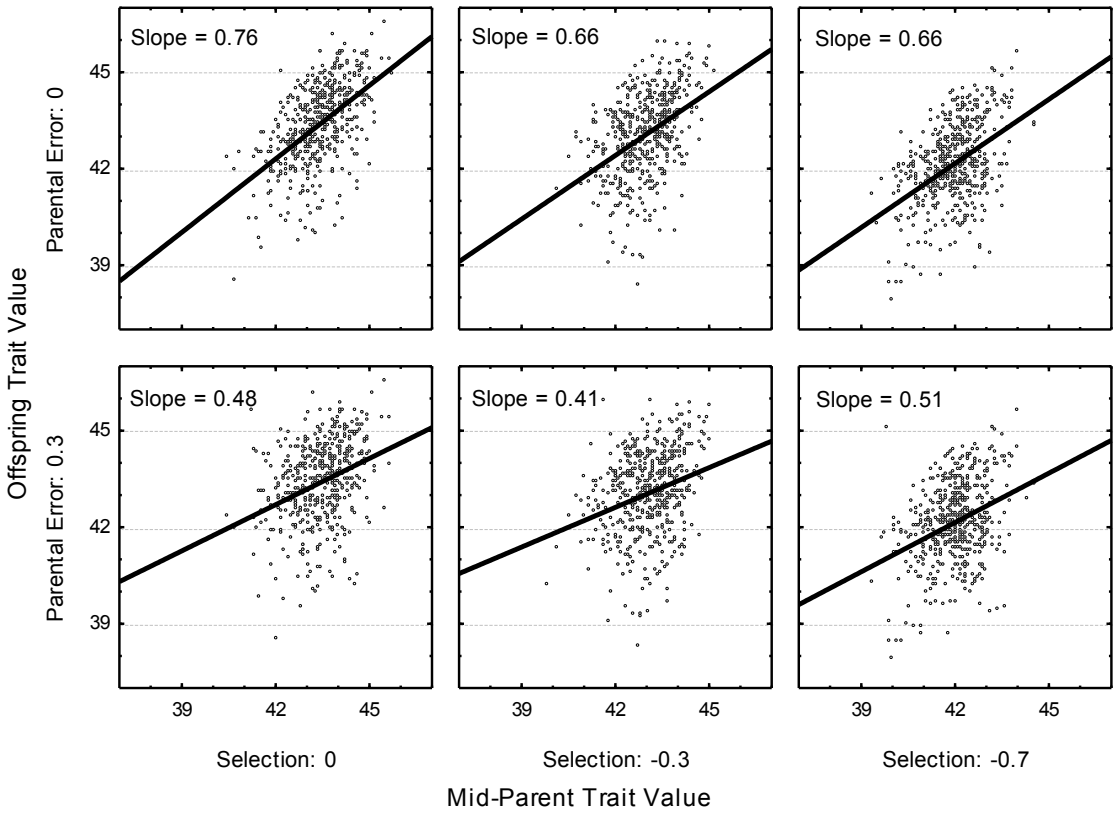


Figure 1.B.1. Linear fit for offspring trait value regressed against mid-parental trait value for the tenth generation one year after birth. Data were simulated for three directional selection strengths (0, -0.3, -0.7) and two parent assignment error rates (0, 0.3) with the following parameter values: cohort size = 1000, parental heritability = 0.7, and annual survival = 0.5. Trait values were drawn from a Weibull distribution with shape and scale parameters equal to 6 and 45 respectively. All parameters besides heritability and selection were held constant. The slope of the linear equation equals the estimate of the trait's heritability. The difference between the slopes for the two error rates for each selection level equals the parental error influence.

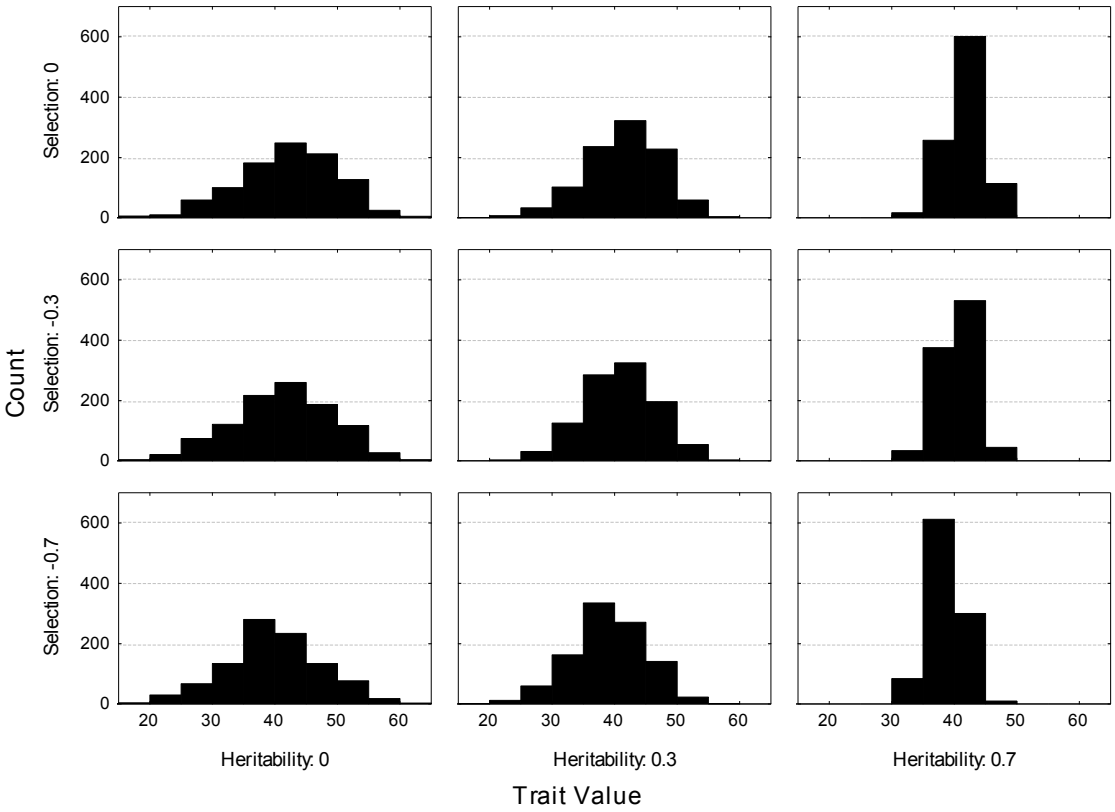


Figure 1.B.2. The affects of heritability and selection on trait value distribution for a single cohort after ten generations. Data were simulated for three heritability levels (0, 0.3, 0.7), and three directional selection strengths (0, -0.3, -0.7) with the following parameter values: cohort size = 1000, and annual survival = 0.5. Trait values were drawn from a Weibull distribution with shape and scale parameters equal to 6 and 45 respectively. All parameters besides heritability and selection were held constant.

PedAgree: Software to quantify error and assess accuracy and congruence for genetically reconstructed pedigree relationships

Abstract

PedAgree is software for rapid comparison of genetically reconstructed pedigrees (RP's). Its two primary functions are 1) to assess accuracy of a RP by comparing it to a known pedigree, and 2) to measure congruence between two RP's. The accuracy function is used to assist in determining confidence for a RP. The congruence function is used to determine the level of agreement between two RP's. This function determines which links within the RP's are identical, and thus more likely to be correct. Congruence assessment between RP's generated by sibship reconstruction (SR) and parentage assignment (PA) programs allows for implementation of the sibship constraint method. This method has been shown to increase assigned parentage accuracy by up to 53%, and to be robust to dataset characteristics that reduce conventional PA accuracies. PedAgree can compare output produced by seven SR and twelve PA programs, and is freely available for download at https://bcrc.bio.umass.edu/pedigreesoftware/.

Program Description

Knowledge of a population's pedigree enables investigation into key questions in evolutionary biology such as natural and sexual selection (Cockburn *et al.* 2008), inbreeding depression (Szulkin & Sheldon 2008), and fitness (Keller *et al.* 2008).

Historically, pedigrees were based upon social interactions and were thus limited to organisms amenable to mating, birth, and parenting observations, typically birds and large ungulates (Kruuk 2004). Recently however, advances in genetic techniques and statistical methods have made it possible to reconstruct a pedigree for virtually any population (Pemberton 2008).

Although a simple concept, the reconstruction of a pedigree holds great potential for error. With social pedigrees, error typically arises from mistaken conclusion of paternity based upon behavioral observations (e.g. O'Connor *et al.* 2006). The use of genetics in pedigree reconstruction brings with it a new suite of error sources, primarily in the form of genotyping error and incomplete population sampling (Wilson & Ferguson 2002). This makes assessment of error effects on reconstruction and assignment accuracies particularly important. Though the effect of pedigree inaccuracies on pedigree-derived metrics is still a largely unexplored area (but see (Charmantier & Reale 2005)), errors in pedigree links can result in heritability and inbreeding estimates that are downwardly biased and less precise (Kruuk 2004; Pemberton 2008), could propagate errors in pedigree-based analysis of fitness (Coulson *et al.* 2006; Pelletier *et al.* 2007), and lead to misinterpretation of dispersal, mating strategy, and reproductive success. It is therefore critical to employ methods that increase reconstructed pedigree (RP) accuracy, and to obtain a measure of accuracy for both RP's and measures derived using them.

PedAgree was developed to assist with accuracy improvement and assessment. It is used for two primary functions: to assess accuracy and quantify error, and to assess output congruence. Accuracy assessment and error quantification is used to assist in obtaining a confidence level for a RP. To do so populations must be simulated with

genetic and demographic attributes characteristic of the true population while recording the population's pedigree. The simulation programs PEDAGOG (Coombs *et al*. 2010a) or EASYPOP (Balloux 2001) are suitable for this purpose. Sibship reconstruction (SR) and/or parentage assignment (PA) analyses are then run on the simulated dataset and compared to the true pedigree using the accuracy assessment function within PedAgree. An accuracy comparison produces two output files, a 'details' file containing information for each individual, and a 'summary' file tabulating the results of the details file.

For SR analyses, the details file records the individual, their true family identifier and size (number of individuals), and their assigned family identifier and size. The summary output file (Figure 1.C.1) displays the number, identifier, and composition of true and assigned full-sib families, followed by the identifier and proportion of assigned families needed to reconstruct each true family. This is followed by the total accuracy of all assigned families which is equal to the total number of individuals minus the minimum number of moves required to convert assigned full-sib family number and composition to true full-sib family number and composition. The file concludes with accuracies of assigned families greater than or equal to a specific size.

For PA analyses, the details file displays the individual, their dam and first assigned parent, their sire and second assigned parent, a description of the comparison assessment for both assigned parents, an individual score, and a relatedness description for incorrectly assigned parents. The comparison assessment falls into one of four categories: right-sampled, right-not sampled, wrong-sampled, and wrong-not sampled. Right and wrong refer to whether the assigned parent matches the true parent. Sampled and not sampled refer to whether the parent was present as a parental candidate, and thus

available for assignment. Score refers to the number of correctly assigned parents and ranges from 0 for both incorrect to 2 for both correct. If a comparison assessment of wrong-sampled is obtained, a relatedness description based on the relationship between the true and assigned parent is recorded and categorized as full-sib, half-sib, or non-sib. If no parent was assigned then the category 'not assigned' is recorded.

The parentage summary file (Figure 1.C.2) begins by listing accuracies for categories based upon the sampling status of the parents. Categories consist of 'neither parent available', 'one parent available', and 'both parents available'. A line tabulating these three categories is also given. Following this is a summary of the relatedness categories for incorrectly assigned individuals when the true parent was available. Tallies are done for situations where only one parent was available for candidacy, and where both parents were available. The next section lists accuracies for all instances where a parent was assigned, followed by accuracies for instances where both of an offspring's parents were assigned. The file concludes by listing assignment numbers and proportions for two classes. The first class contains instances when the true parent was sampled. Results are given for categories of correctly assigned, incorrectly assigned, and incorrectly unassigned. The second class contains instances when a parent was assigned. Results are given for categories of sampled and correctly assigned, sampled and incorrectly assigned, and unsampled and incorrectly assigned.

The second primary function of PedAgree is to compare two RP's and assess their congruence. This function allows for the identification of families or parents that were identically reconstructed or assigned. The RP's are typically generated by different programs, but different runs by the same program can be assessed as well. From analysis

of simulated data, congruent assignments typically possessed higher accuracy (J. Coombs, unpublished data). Thus restricting downstream analyses to the congruent subset should improve accuracy of pedigree derived measures.

There are three categories of comparison based on the type of RP generated: sibship-sibship, parentage-parentage, and sibship-parentage. All three categories work by comparing the second RP to the first. Output files produced by the first two comparison categories are similar to those described for the accuracy comparison, with the exception that relatedness information is not available for PA.

The third comparison category, sibship-parentage, allows for implementation of the sibship constraint (SC) method. This method generates a pedigree output file that assigns the parents with the highest assignment proportions within a full-sib family to all members of that family provided the assignment proportion is greater than or equal to a user-specified value. This should result in higher accuracies for situations where SR accuracy is high and family sizes are large (Wang 2007). Simulations evaluating the robustness of the method and improvement to assigned accuracies support this premise by increasing accuracies by as much as 53% over PA output alone (Coombs 2010). The majority of the increase in accuracy resulted from the removal of assigned parents for instances when true parents weren't sampled, which itself was a function of the proportion of candidate parents sampled.

Currently PedAgree can perform comparisons for RP's generated by eighteen different programs. Accuracy and congruence for SR output can be assessed for COLONY v1.2 (Wang 2004), COLONY v2.0 (Wang & Santure 2009), KINGROUP (Konovalov *et al.* 2004), KINALYZER (Ashley *et al.* 2009), PARENTAGE (Emery *et*

al. 2001), PEDIGREE (Smith *et al.* 2001), and PRT (Almudevar & Field 1999). Accuracy and congruence for PA output can be assessed for CERVUS (Kalinowski *et al.* 2007), COLONY v2.0, FAMOZ (Gerber *et al.* 2003), FAMSPHERE (Carvajal-Rodriguez 2007), GIMLET (Valiere 2002), NEWPAT (Wilmer *et al.* 1999), PAPA (Duchesne *et al.* 2002), PARENTE (Cercueil *et al.* 2002) , PASOS (Duchesne *et al.* 2005), PEDAPP (Almudevar 2007), PROBMAX (Danzmann 1997), and WHICHPARENTS (Hedgecock & Eichert 1999). Additionally, pedigree files created by the SC method can be compared for accuracy and congruence by selecting PEDAGREE as the output software source.

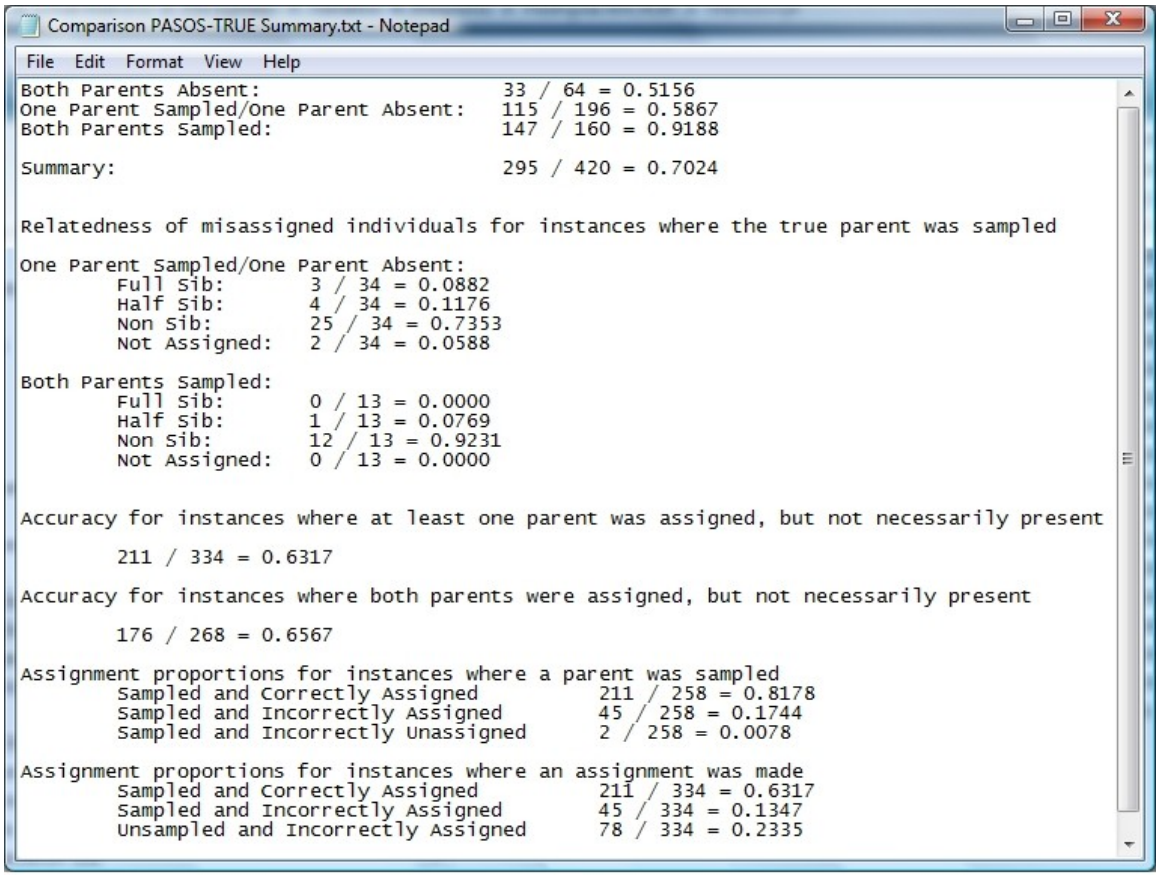
A Windows executable version of PedAgree along with documentation and example data files can be freely downloaded at https://bcrc.bio.umass.edu/pedigreesoftware/. We intend to remain active in implementing new program options as future programs are created or as need arises for existing ones.

Comparison COLONY-TRUE SUMMARY.txt - Notepad

File Edit Format View Help

```
Number of Full Sib Families

True     Assigned
10       9

True Full Sib Families

Family   Size     Member ID's
1)       5        5034     5037     5040     5044     5046
2)       3        5061     5063     5065
3)       3        5082     5083     5085
4)       2        5117     5123
5)       2        5130     5139
6)       1        5051
7)       1        5078
8)       1        5088
9)       1        5108
10)      1        5124


Assigned Full Sib Families

Family   Size     Member ID's
1)       4        5051     5082     5083     5085
2)       3        5061     5063     5065
3)       3        5034     5044     5046
4)       3        5037     5117     5123
5)       2        5078     5108
6)       2        5130     5139
7)       1        5088
8)       1        5040
9)       1        5124


True Family (Size)       Assigned Family (# in True Family/Size)
1 (5)                    3 (3/3)
                         4 (1/3)
                         8 (1/1)
2 (3)                    2 (3/3)
3 (3)                    1 (3/4)
4 (2)                    4 (2/3)
5 (2)                    6 (2/2)
6 (1)                    1 (1/4)
7 (1)                    5 (1/2)
8 (1)                    7 (1/1)
9 (1)                    5 (1/2)
10 (1)                   9 (1/1)


Total    16 / 20 = 0.8000

Accuracy for assigned full sib families of size >= 4
         3 / 4 = 0.7500

Accuracy for assigned full sib families of size >= 3
         11 / 13 = 0.8462

Accuracy for assigned full sib families of size >= 2
         14 / 17 = 0.8235
```

Figure 1.C.1. An example of a sibship reconstruction accuracy summary file.



Comparison PASOS-TRUE Summary.txt - Notepad

File Edit Format View Help

```
Both Parents Absent:                        33 / 64 = 0.5156
One Parent Sampled/One Parent Absent:       115 / 196 = 0.5867
Both Parents Sampled:                       147 / 160 = 0.9188

Summary:                                    295 / 420 = 0.7024


Relatedness of misassigned individuals for instances where the true parent was sampled

One Parent Sampled/One Parent Absent:
        Full Sib:       3 / 34 = 0.0882
        Half Sib:       4 / 34 = 0.1176
        Non Sib:        25 / 34 = 0.7353
        Not Assigned:   2 / 34 = 0.0588

Both Parents Sampled:
        Full Sib:       0 / 13 = 0.0000
        Half Sib:       1 / 13 = 0.0769
        Non Sib:        12 / 13 = 0.9231
        Not Assigned:   0 / 13 = 0.0000


Accuracy for instances where at least one parent was assigned, but not necessarily present

        211 / 334 = 0.6317

Accuracy for instances where both parents were assigned, but not necessarily present

        176 / 268 = 0.6567

Assignment proportions for instances where a parent was sampled
        Sampled and Correctly Assigned          211 / 258 = 0.8178
        Sampled and Incorrectly Assigned        45 / 258 = 0.1744
        Sampled and Incorrectly Unassigned      2 / 258 = 0.0078

Assignment proportions for instances where an assignment was made
        Sampled and Correctly Assigned          211 / 334 = 0.6317
        Sampled and Incorrectly Assigned        45 / 334 = 0.1347
        Unsampled and Incorrectly Assigned      78 / 334 = 0.2335
```

Figure 1.C.2. An example of a parentage assignment accuracy summary file.

CHAPTER II

GENETICALLY RECONSTRUCTED PEDIGREES: THE COSTS AND BENEFITS OF USING FULL-SIBLING STRUCTURE TO CONSTRAIN PARENTAGE ASSIGNMENTS

Abstract

We present a simple yet effective method to improve accuracy of parentage assignments by an average of 47% compared to assignments made using the parentage assignment (PA) programs PEDAPP (39%), PASOS (53%), and CERVUS (50%) as measured over a wide range of simulated scenarios. The method, termed sibship constraint (SC), uses the results of sibship reconstruction (SR) performed on a cohort to constrain assignments from PA output. It works by assigning the PA candidates allocated to the greatest proportion of offspring within a reconstructed full-sibling family to all members of that family. A user-specified minimum threshold value determines which candidate(s) to keep based on assignment proportions. Comparisons were made between output produced by the SC method and PA programs for four measures of accuracy evaluated for the following eight variables: minimum threshold value, SR program used, mating strategy, mean family size, proportion of true parents sampled, number of loci used, genotyping error rate, and cohort assignment error rate. The cost of using the SC method was a decrease in assignments made to offspring whose true parents were sampled by 9% compared to PEDAPP and PASOS, and 21% compared to CERVUS

outputs. However, this cost was more than offset by the benefit of a decreased number of assignments made to offspring whose true parents were not sampled by 80% (PEDAPP), 82% (PASOS), and 84% (CERVUS), which resulted in marked improvement to assigned accuracies. The SC method is highly flexible in that it can use outputs from six SR and twelve PA programs, with all SR-PA pairings possible. Additionally, the method is fully automated within the freely-available software program PEDAGREE.

Introduction

Knowledge of a population's pedigree enables investigation and insight into numerous evolutionary, ecological, and behavioral processes that would otherwise be unattainable (Kruuk & Hill 2008; Wilson & Ferguson 2002). Examples of such processes include dispersal (Saenz-Agudelo *et al*. 2009; Szulkin & Sheldon 2008), mating strategy (Theriault *et al*. 2007), reproductive success (Jones *et al*. 2007; Taggart *et al*. 2001), sexual selection (Grant & Grant 2008), natural selection (Garant *et al*. 2004), trait heritability (Kruuk *et al*. 2002), and speciation (Svedin *et al*. 2008).

Even though the potential utility of pedigrees has been known to geneticists for over a century, application of pedigrees to studies involving naturally reproducing populations has been limited (Pemberton 2008). A partial explanation for this is that many species display reproductive and parental behaviors that make it extremely difficult or even impossible to determine parentage from social interactions alone. Thus, creation of pedigrees for these populations has depended on the discovery of appropriate genetic markers coupled with the development of relationship reconstruction algorithms (Blouin

2003; Jones & Ardren 2003). Over the last decade, development of more widely applicable reconstruction algorithms coupled with greater availability of informative markers and decreased cost of molecular techniques has resulted in a surge of studies using genetically reconstructed, multi-generational pedigrees to address critical eco-evolutionary questions using wild populations (Pemberton 2008).

Overall accuracy of reconstructed pedigrees is key for correct interpretation of downstream analyses that depend on pedigrees (Morrissey *et al*. 2007). Although not extensively investigated, initial studies on effects of pedigree errors have reported downward bias in measures for both trait heritabilities (Charmantier & Reale 2005) and inbreeding depression (Pemberton 2008). Additionally, erroneous links could lead to incorrect inferences regarding dispersal, mating strategy, reproductive success, and sexual selection.

In genetically reconstructed pedigrees, incorrect links arise from an inability of the parentage assignment (PA) algorithm to adequately resolve relationships. This occurs primarily when the set of genetic markers has reduced exclusion probability (Gerber *et al*. 2000), but can also be affected by genotyping errors and mutations (O'Reilly *et al*. 1998), and by incomplete sampling of parental candidates (Wilson & Ferguson 2002). Additionally, most PA algorithms evaluate potential parents for one offspring at a time (Wang 2007). This increases the probability of assignment error, especially when dataset quality is reduced, because a single offspring provides information for only half of the alleles in the parental genotype, thus not making full use of the genetic information.

In contrast, most sibship reconstruction (SR) algorithms assess the likelihood of offspring partitions for the sample as a whole (Smith *et al*. 2001; Wang 2004). Sieberts *et*

al (2002) and Wang (2007) both demonstrated that the power to infer relationships increases dramatically with simultaneous analysis of multiple individuals. For example, sibship exclusion only becomes possible with analysis of at least three individuals since it is possible for two full-siblings to not share any alleles for a set of codominant markers. Thus, larger full-sibling partitions possess greater exclusionary power, and are therefore more reliable than smaller full-sibling partitions (Wang 2007).

Given this, we propose a method that utilizes otherwise ignored family structure within a sample to improve assignments made by pair-wise PA algorithms. The method uses the results of SR to evaluate the agreement of parentage assignments. For a full-sibling family, the proportion of assignments made for each parental candidate out of all potential assignments is quantified. For example, a full-sibling family of size ten would have a total of twenty parental assignments. If parental candidate A was assigned to eight offspring and parental candidate B was assigned to five offspring, then A would have an assignment proportion of 0.4 (8/20) (where 0.5 is the maximum proportion possible) and B would have a proportion of 0.25 (5/20). The method then uses a user-specified minimum threshold value to determine whether to discard the top parental candidates, or assign them to all offspring within the family. In the previous example, a specified minimum threshold value of 0.2501 would assign candidate A to all offspring and discard candidate B. This process is then repeated for all full-sibling families of size greater than or equal to a user-specified value.

We evaluated the performance and robustness of the method, henceforth referred to as sibship constraint (SC), using simulated datasets. A total of eight variables were investigated, and results produced by the SC method were compared to those produced

by the PA programs alone to assess accuracy and assignment rates. The results illustrate the limitations of the method, and identify the costs and benefits of SC compared to traditional PA programs.

Materials and Methods

Simulations

Data were simulated using the program PEDAGOG v1.2 (Coombs *et al*. 2010a) because of its ability to track individual pedigree and genotype information, allow for manipulation of genetic, demographic, and error parameters, and automatically format the simulation output into input files for pedigree reconstruction programs. The baseline population was parameterized to have five age classes and a constant cohort size of 500 animals with a 0.5 probability of being female. Founding cohorts for the population were drawn from a population pool of 10,000 animals whose genotypes were assigned randomly from allele frequencies for the set of eight primary loci (Table 2.1). Allele frequencies for loci 1-7 (King 2003) and locus 8 (King *et al*. 2005) were derived from a brook trout (*Salvelinus fontinalis*) population located in the Fridley Gap watershed in West Virginia (M. Hudy, unpublished data). The allele frequencies for the loci are shown in Appendix A.

Subsequent cohorts reproduced using a polygamous mating system where all animals age one or older matured annually, mate number was drawn from a Poisson distribution with mean and standard deviation of two, and males and females within three

generations of each other were allowed to mate. Fecundity was drawn from a gamma distribution with shape and scale parameters both equal to three. For females mating with multiple males, the proportion of offspring sired was size-dominant with the largest male's proportion drawn from a normal distribution with mean of 0.8 and standard deviation of 0.05. Non-dominant males were assigned a randomly generated proportion of the remaining offspring. Annual survival probabilities for age classes zero through four were 0.41, 0.66, 0.81, 0.90, and 0.95, and were the same regardless of sex or maturity status. The population was sampled after ten generations with capture probabilities of 0.9 for age class zero animals, and 0.95 for the remaining age classes. Sex of captured individuals was unknown. All simulations were replicated ten times.

A total of six variables were altered from the baseline population model to evaluate their effects on accuracy of the SC method. The six variables were grouped into categories of intrinsic population characteristics, power to perform pedigree reconstruction, and error effects. Variables associated with intrinsic population characteristics were mating strategy and female fecundity. Mating strategy simulations were run for both monogamous and polyandrous scenarios in addition to the polygamous baseline scenario. Monogamy restricted both males and females to only one mate per reproductive season, but did not constrain individual pairs to mate for life. Polyandry set male mate number to one per mating season while the number of female mates was drawn from a Poisson distribution with mean and standard deviation of two. Female fecundity was adjusted to result in one scenario of lower than baseline fecundity, and two scenarios of higher than baseline fecundity. Fecundities for these scenarios were drawn

from gamma distributions with shape parameters of 1.75 (lower), 4.5 (higher), and 6.0 (highest), and scale parameters all equal to three.

The variables associated with pedigree reconstruction power were the proportion of true parents sampled and the number of loci used for reconstruction analyses. The baseline scenario resulted in a sampling of approximately 60% of true parents. To produce true parent capture levels of 20%, 40%, 80%, and 98%, either capture probability was adjusted, a sampling event was added during the ninth generation, or a combination of both were used. To evaluate the effect of altering the number of loci, a set of eight supplemental loci were added to the set of eight primary loci (Table 2.1). The supplemental loci were randomly generated in PEDAGOG with allele number and allele frequency restrictions forcing them to be similar to those of the primary locus-set (mean expected heterozygosities equal to 0.79 (primary), and 0.81 (supplemental)). Simulations were run using all sixteen loci, and the program CREATE v1.2 (Coombs *et al*. 2008) was used to make additional SR and PA input files for the first four, eight, and twelve loci from Table 2.1.

Variables evaluating the effect of error on the accuracy of the SC method involved increasing genotyping error and cohort misclassification rates from their baseline values of zero. Simulations were conducted using 0.01, 0.03, and 0.05 locus-specific genotyping error probabilities, and 0.05, 0.10, and 0.15 cohort misclassification probabilities. Genotyping error events consisted of miscalling the true allele as either an adjacent allele, or a random allele, both at a probability of 0.5. A cohort misclassification event assigned the animal to an older cohort if the animal's length was greater than or equal to the mean of the population, or a younger cohort otherwise.

Pedigree Reconstruction

Both SR and PA were performed for each replicate of all simulated scenarios. SR was conducted using the software programs COLONY v1.2 (Wang 2004), KINGROUP v2_090306 (Konovalov *et al*. 2004), and PEDIGREE v2.0 (Smith *et al*. 2001). The following program settings were used as they consistently led to the highest accuracies. For COLONY, the genotyping error rate was set to 0.005 when the simulated error probability was zero, or to the simulated probability if greater than zero. For KINGROUP, the descending ratio full-sibling reconstruction algorithm was used with the primary hypothesis set to full-siblings and the null hypothesis set to half-siblings. For PEDIGREE, a control file containing four runs was used. For each run the number of iterations was set to five million, the full-sib constraint was set to one, the weight was set to one, and the seed was set to negative one. For the four runs, the temperature was set to 5 (Run 1), 15 (Run 2), 25 (Run 3), and 35 (Run 4). The sibship reconstruction from the partition returning the highest score was used for analyses.

PA was performed using the programs PEDAPP v1.1 (Almudevar 2007), PASOS v1.0 (Duchesne *et al*. 2005), and CERVUS v3.0.3 (Kalinowski *et al*. 2007; Marshall *et al*. 1998). These three programs were chosen because they allow for both parents to be unknown, parent sexes to be unknown, and incomplete parental sampling, all of which are likely to occur when working with wild populations. The following program settings were used as they consistently led to the highest accuracies, and were representative of

the population dynamics that generated the simulated samples. For PEDAPP, the permissible parent-offspring age difference was set to greater than or equal to one and less than or equal to five, and the likelihood derived pedigree was used for analyses. Because PEDAPP analyzes all cohorts simultaneously, the 'cohort clipper' option of the software program 3-In-1 (J. Coombs, Available for download at https://bcrc.bio.umass.edu/pedigreesoftware/) was used to extricate assignments for just the cohort of interest. For PASOS, the non-sexed allocation option was always used, and the maximum offset tolerance was set to zero for all analyses, including scenarios involving changes in genotyping error rate. For CERVUS, internal simulations were run for each scenario to establish delta values for assigned confidences. For each internal simulation the number of offspring simulated was set to 10,000, the number of candidate parents was set to 575, and the confidence levels were set to 50% (relaxed) and 90% (strict). A relaxed value of 50% was selected to increase the number of assignments made by CERVUS. Outputs using this confidence level were used for all SC analyses. A strict confidence level of 90% was selected to provide a comparison of accuracy and proportion of parents assigned between this commonly used level of CERVUS and the SC method using all three PA programs. The internal simulation parameter for proportion of candidate parents sampled was set to the mean of the ten replicates for each simulation scenario which was acquired from the PEDAGOG output by using the 'mates and candidates' option of the software program 3-IN-1. The proportion of loci mistyped parameter was set to 0.005 for all scenarios except those altering genotyping error rates for which the value was set to the PEDAGOG simulation probability.

Analyses

All parent assignments produced by the SC method were created using the reconstructed-reconstructed option in the program PEDAGREE v1.04 (Coombs *et al* 2010b). This option compares two output files created by SR and PA programs. Initial analyses evaluated the effects of two analysis parameters: the SR program used, and the minimum threshold value. The minimum threshold value determines whether a candidate is assigned to all members of the full-sibling family based on whether the candidate's proportion of assignments within the family is greater than or equal to the specified value. Effects of both of these variables were analyzed only for the baseline simulations and using full-sibling families of size two or greater (Tables 3 and 4). Based upon these results, analyses for all remaining scenarios used COLONY as the source of SR output and a minimum threshold value equal to 0.2501, while continuing to restrict full-sibling family size to greater than or equal to two.

Means and 95% confidence intervals were calculated from the ten replicates for each scenario for total accuracy, assigned accuracy, the proportion of correct assignments when the true parent was sampled, and the proportion of incorrect assignments when the true parent was not sampled (Table 2.2). Total accuracy (TA) assessed the correctness of assignments for all offspring, including instances when no assignment was made. Assigned accuracy (AA) only evaluated the correctness of instances when a parent was assigned. True parent sampled and correctly assigned (SA) represented the proportion of assigned parents that were correctly made for instances where the true parent was sampled. True parent not sampled and incorrectly assigned (NI) represented the

proportion of assigned parents that were incorrectly made for instances when the true parent was not sampled, referred to by Duchesne *et al*. (2005) as over-allocation. Only assignments for offspring belonging to reconstructed families of size two or greater were used in accuracy calculations. This proportion was 85% (±1%) of the entire cohort when averaged over all scenarios.

Accuracy values were calculated for output from each PA program by itself, and for output produced by the SC method using that PA program's output as the data source. Additionally, AA and SA values were calculated for CERVUS output acquired using the strict 90% confidence level setting. All calculations were made using the reconstructed-true option in PEDAGREE, and using the 'true genotypes' output file from the associated PEDAGOG simulation as the reference to the true population pedigree.

Results

The mean number of offspring used for SR analyses was 186 (±2), and the mean number of candidates used for PA analyses was 375 (±48). Within a cohort, the mean number of full-sibling families was 71 (±6), the mean family size was 2.7 (±0.3), and the mean largest full-sibling family size was 9.9 (±0.8). Additional population attributes along with accuracies for the raw SR and PA outputs are available by request from the author.

Percent differences for accuracy measures between SC and PA outputs are shown in Table 2.3. The SC method consistently produced higher TA and AA values compared to those produced by PA programs alone. SA values produced using the SC method were

reduced compared to those from associated PA output. However, these reductions were an order of magnitude lower than the reductions in NI values for PEDAPP and PASOS, and one-fourth that of CERVUS (Table 2.3). Thus the SC method was more conservative in its assignment of parents than the PA programs which resulted in a slight reduction of assignment to offspring whose true parents were sampled, but a drastic reduction in assignments to offspring whose true parents were not sampled.

The SC method also produced greater accuracies than output from CERVUS acquired using a 90% confidence level (Table 2.3). For AA values, output using the SC method was on average 7.5% higher than 90% CERVUS output. Of even greater significance was that SA values from output using the SC method were 125% to 353% higher than those from 90% CERVUS output. Thus the SC method assigned significantly more parents than 90% CERVUS, and did so with greater accuracy.

Of the three PA programs used as data sources for SC, output using PEDAPP resulted in slightly improved accuracies over output using PASOS, while use of CERVUS as the output data source produced the poorest results (Tables 4-7). Overall, SC accuracy values averaged from Tables 4 through 7 for PEDAPP, PASOS, and CERVUS were: TA – (0.84, 0.83, 0.68), AA – (0.92, 0.91, 0.93), SA – (0.75, 0.73, 0.45), and NI – (0.04, 0.05, 0.05). Thus CERVUS was significantly more conservative in its assignments compared to the other two programs.

Analysis Variables

COLONY consistently reconstructed full-sibling families with the highest accuracies (data shown in supplementary material). This in turn produced the highest TA, AA, and SA values and the lowest NI values for SC output when COLONY was used as the SR data source (Table 2.4). PEDIGREE output resulted in the second highest accuracy values followed by KINGROUP. Use of output from either these programs as the SR data source for the SC method resulted in substantial decreases to SA values (Table 2.4).

The minimum threshold value dictated the conservativeness of assignments made using the SC method. Alteration of the minimum threshold value produced a trade-off among SA, AA, and NI values. Smaller minimum threshold values resulted in higher SA values, but lower AA and higher NI values. Larger minimum threshold values resulted in higher AA and lower NI values, but lower SA values. A minimum threshold value of 0.2501 resulted in the highest AA and lowest NI values, while maximizing the SA value (Table 2.4).

Population Variables

Mating strategy had surprisingly little effect on SC output accuracies (Table 2.5). There was a slight trend towards decreased SA values when progressing from monogamous to polyandrous to polygamous mating systems. However, AA and NI values remained essentially constant over the same progression

Alternatively, family size had a pronounced effect on the number of assignments made for SC output. An increase in average family size resulted in increased SA and TA values (Table 2.5). Mean family sizes of 3.5 and 4.3 resulted in the only instances when SA values produced by the SC method were higher than those produced by PA programs alone. This trend occurred for both PEDAPP and PASOS output.

Power Variables

The proportion of true parents sampled had substantial effects on both the number and accuracy of parent assignments made by the SC method, particularly when PEDAPP or PASOS output was used as the PA data source (Table 2.6). NI values at low sampled parent proportions (0.2 and 0.4) were significantly elevated compared to values when sampled parent proportions were greater than or equal to 0.6. This result in turn produced the opposite pattern for AA values, where accuracy decreased as sampled parent proportion decreased. Comparatively, CERVUS' more conservative output resulted in decreased SA values instead of increased NI values as the proportion of sampled parents decreased. This pattern ultimately resulted in increased AA values at low sampled parent proportions. There were no clear linear trends in SA values for SC output using either PEDAPP or PASOS as data sources.

Increasing the number of loci used for analyses resulted in an appreciable change in the number of assignments made for instances when true parents were sampled, but only a slight change to the accuracy of made assignments (Table 2.6). For SC output using PEDAPP and PASOS data sources, SA values increased approximately 21% as the

number of loci used increased from eight to twelve, and 26% for an increase from eight to sixteen. Comparatively, AA values only increased 4% and 5% for the same increased in number of loci used.

Error Variables

Both genotyping and cohort misclassification errors resulted in the SC method making fewer assignments (decreased SA values) with increased rates of error (Table 2.7). However, AA values remained stable as error rates increased. Thus, the SC method maintained assignment accuracies when faced with dataset degradation by sacrificing the number of assignments made.

Discussion

The SC method uses SR output to delineate full-sibling families and then assign the most commonly allocated parental candidates from PA output to the entire family provided a candidate's proportion of assignments exceeds a user-specified minimum threshold value. Accuracies produced using the SC method were compared to those produced using traditional PA programs for simulated datasets investigating the effects of eight variables. The SC method produced substantially higher TA and AA values while simultaneously reducing NI values compared to PA output alone (Table 2.3). The one cost of the SC method was a decrease in SA values compared to PA output. However, for the current simulations this cost was only 11% to 24% of the benefit gained from

reducing NI values depending upon which PA program was used (Table 2.3). Thus, use of the SC method resulted in more accurate assignments than use of PA programs alone.

Of the three PA programs used for this study, PEDAPP consistently returned the highest TA and AA values for raw output (shown in supplementary material). This subsequently led to the highest TA, AA, and SA values and the lowest NI values when its output was used by the SC method (Tables 4-7). Use of PASOS output for the SC method resulted in only slightly reduced accuracy values despite its raw output consistently having lower AA and higher NI values (Tables 4-7).

The reason that nearly identical accuracies were produced by the SC method when using PA sources with, in some instances, significantly different accuracies stems from the mechanism behind incorrect assignments. Incorrect parents were assigned more frequently when fewer loci were used in reconstruction analyses and true sampled parent proportions were reduced (Table 2.6). These conditions provided assignment algorithms greater opportunity to assign a single false parent to offspring with at least one unsampled parent because only one of the candidate parent's alleles had to match either of the offspring's alleles at each locus. For the baseline simulation scenario (60% of true parents sampled, eight loci used for analyses) TA values for offspring with zero, one, and two true parents sampled were 57%, 74%, and 92% for PEDAPP, and 54%, 60%, and 91% for PASOS (data not shown).

Additionally, likelihood methods within PA programs rank all possible parents based upon the alleles in the candidate's genotype versus allele frequencies in the population (Jones & Ardren 2003; Marshall *et al*. 1998). Thus for full-sibling families with one or more unsampled parents, different false parents could be assigned to different

offspring within the same full-sibling family because of differing offspring genotypes. Indeed, for the baseline simulation scenario, the number of unique candidates assigned to each parent of a full-sibling family where a minimum of two assignments were made were 3.0 (±0.36), 2.0 (±0.12), and 1.2 (±0.06) for families with zero, one, and two true parents sampled (data not shown).

As the number of candidates assigned to a full-sibling family increases, each candidate's proportion of assignments decreases. Thus, the use of an adequate minimum threshold value for the SC method is an effective means of ensuring that full-sibling families assigned multiple parents do not retain any of them (Table 2.4). Results from the SC method supported this fact by producing TA values for offspring with zero, one, and two parents of 89%, 81%, and 87% for COLONY-PEDAPP output, and 87%, 77%, and 85% for COLONY-PASOS output for the baseline scenario (data not shown). Reduction in accuracy for instances when both parents were sampled was caused by reductions in SA values, not increases in NI values (data not shown).

Results differed in important ways among PA programs. Compared to PEDAPP and PASOS, CERVUS resulted in similar AA values, but significantly lower SA values for both PA and SC output (Tables 4-7). One reason for this was CERVUS' use of internal confidence levels to classify assignments (Marshall *et al*. 1998). This internal filter resulted in a reduced number of assigned parents even with the relaxed confidence level set to 50%. The reduced number of assignments was also a function of the proportion of sampled parents. For sampled parent proportions of 60% or less, available assignment proportions from CERVUS output declined at a much faster rate than for output from either PEDAPP or PASOS (Table 2.6). The reason for this pattern is that

CERVUS uses a parent-pair assignment algorithm when trying to assign both parents (CERVUS user manual). Thus, failure to sample one of the parent pair renders correct assignment impossible (Jones & Ardren 2003). However this pattern seems to apply to CERVUS in general, because both Marshall *et al* (1998) and Wilson and Ferguson (2002) reported similar results when conducting paternity analyses.

This trend of inferior performance by CERVUS when the proportion of sampled candidate parents is reduced is of concern for two reasons. The first is that there has been a dramatic increase in the use of multi-generational pedigrees in studies involving natural populations over the last decade (Pemberton 2008). Given population processes and sampling logistics, the vast majority of those studies almost assuredly contained incomplete sampling of candidate parents. The second reason is that CERVUS is the PA program used most often. From March of 2007 (the date of the most recent publication for the three PA programs used in this study) until July of 2009, the numbers of papers citing each program were 3 (PEDAPP), 9 (PASOS), and 552 (CERVUS) (Web of Science). This suggests that many studies may have analyzed reduced pedigrees stemming from the conservative nature of CERVUS assignments.

This study quantified the reduction in pedigree information by comparing SA values acquired using the SC method to those produced using CERVUS with a 90% confidence level. The net result was an increase in SA values by 353% (±135%) if PEDAPP was used as the SC data source, 346% (±136%) if PASOS was used as the SC data source, and 125% (±25%) if CERVUS with a 50% confidence level was used as the SC data source (Table 2.3). Thus for this study, downstream analyses using pedigrees produced by the SC method would have access to three and half times the number of

pedigree links compared to one produced using CERVUS with a 90% confidence level. Furthermore, AA values from the SC method were approximately 7.5% greater than those produced by CERVUS with a 90% confidence level output (Table 2.3). Thus, not only were more assignments made using the SC method, but those assignments were made with greater accuracy.

AA values produced by the SC method proved to be remarkably robust to changes in parameter values. Of the twenty-five unique scenarios produced by altering a value for one of the eight variables investigated, only two resulted in mean AA values of less than 80% (loci used = 4, threshold value = 0), and only four resulted in mean AA values of between 80% and 90% (threshold value = 0.1667, sibship program = KINGROUP, sampled parents = 0.2 and 0.4). The remaining nineteen variables all resulted in assigned accuracies greater than 90% (Tables 4-7) indicating that SC can provide accurate assignments across a wide range of parameter values.

While the accuracy of the SC method was robust to parameter variation, different parameter values resulted in changes to SA values. Parameter values resulting in increased SA values while maintaining AA levels included simplification of mating strategy (Table 2.4), increased family size (Table 2.4), and increased number of loci used (Table 2.5). Mechanisms behind these trends operated in two different ways. An increase in the number of loci increased the amount of information available to resolve putative relationships (Wang & Santure 2009). The end result was an increase in both the number and accuracy of parentage assignments (Table 2.6), and an increase in SR accuracy (data not shown). This in turn produced a greater number of retained correct parents when employing the SC method because assigned parent proportions exceeded the minimum

threshold value with greater frequency and were assigned to offspring within more accurate sibship partitions (Table 2.6).

The second mechanism providing increased SA values resulted from increases in mean full-sibling family size which in turn improved SR accuracy. Wang (2007) demonstrated that for a given set of markers, inferred families become increasingly reliable with increased size regardless of methodology used in reconstruction. This was evidenced in this study as COLONY SR accuracy increased 14.3% as mean family size was deliberately increased from 1.9 to 4.3 (data not shown). Likewise, when mating strategy was simplified to monogamy from polygamy, mean family size indirectly increased from 2.5 to 3.5 resulting in a 2.7% improvement to SR accuracy (data not shown).

In contrast, increased genotyping and demographic error rates resulted in decreased SA values (Table 2.7). Increased demographic error rates reduced SA values through decreased mean family sizes (data not shown). Full-sibling family sizes were reduced by removal of misclassified individuals, while the overall number of full-sibling families was increased by addition of misclassified individuals from other cohorts. A demographic error rate of 15% reduced mean family size from 2.5 (±0.14) to 2.0 (±0.07) (data not shown) resulting in a 12% reduction in SA values for COLONY-PEDAPP output (Table 2.7).

Genotyping errors had a pronounced effect upon SA values for output produced using the SC method (Table 2.7). SA values from SC output averaged for all three PA data sources decreased by 10%, 20%, and 38% as genotyping error probability increased from 1% to 3% to 5% (Table 2.7). The primary reason for this was that SA values

produced by PA output were also reduced by an increased genotyping error rate (Table 2.7). Given that AA and NI values from PA output remained approximately constant as genotyping error rate increased (Table 2.7), the reduction in SA values must be explained by a decreased number of assignments. Indeed, comparison of PEDAPP output run on true and error containing versions of simulated datasets for genotyping error probabilities of 1%, 3% and 5% resulted in reductions in the number of assignments made by 3%, 13%, and 16% (data not shown).

SR programs proved to be less susceptible to genotyping errors than PA programs (data not shown). Comparison between true and error containing versions of datasets for 5% genotyping error simulations reduced SR accuracies by 5.7% (COLONY), 5.0% (KINGROUP), and 11.3% (PEDIGREE) (data not shown). Genotyping errors split affected individuals from their true full-sibling families. Comparison of number of families between true and error containing versions of datasets resulted in average changes of -2.4 (COLONY), 5.5 (KINGROUP), and 9.3 (PEDIGREE). We believe the negative value for COLONY to be a function of its internal error-handling capability (Wang 2004) which made the joining of two single individuals together more likely. Comparatively, the number of families for KINGROUP and PEDIGREE, which do not have error-handling capabilities, both increased in the presence of errors. The net result is a decrease in size for larger full-sibling families which in turn results in decreased SA values (Table 2.4).

Of the three sibship programs investigated, COLONY consistently returned the highest SR accuracies, followed by PEDIGREE, and then KINGROUP (Table 2.4 and unpublished data). In addition to its error-handling capabilities, COLONY also dealt with

the effects of polygamy better. For simulations specifying a monogamous mating system, COLONY and PEDIGREE had identical SR accuracies (94%), but a change to polygamy resulted in a 6% difference in SR accuracies between the two programs (92% vs. 86%) (data not shown).

SR accuracies for KINGROUP were usually well below those of COLONY and PEDIGREE. Only when the number of loci used was increased to twelve or sixteen did accuracies begin to approach those of the other two programs (data not shown). This suggests that the descending ratio method within KINGROUP requires a lot of genetic information to perform well.

Additional advantages to using the SC method can be classified into areas of flexibility and resource conservation. Flexibility is present in three areas. The first two are user-specified options. One enables the user to specify the minimum full-sibling family size for which the SC method should be applied. The second allows the user to specify a minimum threshold value to determine when to keep a parent for a full-sibling family. These options allow the user to be more or less conservative depending upon the situation. For example, the simulations with sampled parent proportions of 0.2 and 0.4 returned relatively low AA and high NI values when a minimum threshold value of 0.2501 was used (Table 2.6). Raising the minimum threshold value to 0.3334 resulted in improved mean AA and NI values (data not shown). For simulations with sampled parent proportion equal to 0.2, AA values increased from 82% to 95%, and NI values decreased from 0.16 to 0.05. For simulations with sampled parent proportion equal to 0.4, AA values increased from 86% to 95%, and NI values decreased from 0.10 to 0.04 (data not shown).

These options also enable the user to perform the SC method multiple times using different combinations of minimum threshold value and full-sibling family size. For example, a value of 0.2501 may be specified for families greater than or equal to two, and a value of 0.1667 for families greater than or equal to three. Assignments from the two pedigrees could then be combined to create the final pedigree. Determination of what combinations of sibship size and minimum threshold value return acceptable accuracies must be determined through simulation and recovery analyses with a simulation program such as PEDAGOG.

The other area the SC method provides flexibility in is its capacity to use sibship and parentage data from several different sources. This is possible because the method is fully implemented within the software program PEDAGREE which has the ability to read in data from six SR programs and twelve PA programs. This allows for output from any combination of SR and PA programs to be selected and used for SC. What programs are used depends upon the quality of the dataset, the mating strategy and family structure within the cohort, and the subsequent use of the pedigree. For simulations conducted for this study the SR program COLONY and the PA program PEDAPP consistently returned the highest accuracies (Tables 4-7). However, for populations with high rates of monogamy and/or large full-sibling families the SR PEDIGREE returned similar accuracies as COLONY in about a third of the computation time.

Advantages in the area of resource conservation deal with reduced expense, and reduced data acquisition and computation times. Reduced expense and data acquisition time emerge from the need to use fewer loci to achieve similarly high AA values (Table 2.6). For the current study, AA values produced by the SC method using eight loci were

only slightly lower than AA values for PA output that used sixteen loci (Table 2.6). Granted, there was a cost in the form of reduced SA values with fewer loci, but in many situations this may be acceptable. Alternatively, the SC method's ability to return high AA values using a reduced number of markers, and thus a reduced amount of human and machine time required for data acquisition, may allow studies with financial and/or logistic restrictions to be conducted. Additionally, a reduced number of loci would result in a dataset with fewer genotyping errors (Jones & Ardren 2003; O'Reilly *et al*. 1998). This would effectively offset a portion of the cost in reduced SA values since genotyping errors also decrease this measure (Table 2.7). Also, Wang (2004) reported that using an increased number of loci with genetic errors can result in worse relationship estimates if they are ignored.

The SC method can also conserve resources by reducing computing time. If SR and PA output already exist for the cohort of interest, the SC method can be performed in a few seconds. If not, both will have to be run with time to completion varying among SR and PA programs. Comparative run times for the programs used in this study for baseline scenario conditions (≈181 offspring, 379 parents, 0.6 candidate sampling) using a 2.0 GHz CPU were as follows: KINGROUP, PEDAPP, and PASOS completed in seconds, PEDIGREE and CERVUS (not counting simulations) finished in less than five minutes, and COLONY v1.2 concluded in approximately fifteen minutes.

Comparatively, COLONY v2.0 (Wang & Santure 2009), a recently released update to COLONY v1.2 that has the capability to perform SR and PA jointly, can require weeks or longer to reach completion. Run time in COLONY v2.0 is dependent upon the number of offspring, the parental mating strategies, the number and

informativeness of loci, and whether genotyping errors are present (Wang & Santure 2009). For datasets where one or both parents are polygamous and genotyping error rates are greater than zero, increasing the number of offspring increases the number of possible partitions at a greater than exponential rate (COLONY v2.0 User Manual). Additionally, the authors suggest performing multiple runs of increasing length until an acceptable level of data convergence is reached, and then only using congruent data for downstream analyses.

For the first two repetitions of the baseline scenario, runtime using COLONY v2.0 was moderate, requiring approximately two hours for a single run of medium length. However, the first two repetitions of the 3% genotyping error scenario increased the time for a single medium run to sixteen hours while producing no change in runtime for the other programs. Analysis of a dataset from a wild brook trout population with a polygamous mating system, 476 offspring, 388 candidate fathers, 308 candidate mothers, and twelve loci each with an estimated 1% genotyping error rate had been running for four weeks at the time of manuscript submittal and still not reached convergence. Comparatively, the SC method using PEDIGREE output as the SR data source and PEDAPP output as the PA data source required only fifteen minutes for all programs to reach completion.

Nevertheless, the cost of increased analysis time may be worthwhile depending upon the structure and size of the dataset and questions for which the pedigree will be used. Output from COLONY v2.0 for the first two repetitions of the baseline scenario resulted in an increase in AA value of 1%, and an increase in SA value of 10% over SC output using COLONY v1.2 and PEDAPP data sources (unpublished data). However, the

first two repetitions of the 3% genotyping error resulted in a decrease in AA value of 3% but an increase in SA value of 33% (unpublished data). Further investigations are needed to assess the costs and benefits for the two methods.

In summary, the use of SC for PA has numerous advantages over traditional pair-wise PA programs alone. The method is best suited for populations with significant family structure as it requires sampling and accurate partition of multiple full-sibling family members to improve upon accuracies of the PA output. It particularly outperformed traditional PA programs when proportions of true parents sampled were reduced by removal of over-allocated parent assignments. The SC method also significantly outperformed the leading PA program CERVUS. Output produced by the SC method increased SA values by up to 350% and AA value by up to 9% compared to CERVUS output acquired using a 90% confidence level. Overall, the SC method has the ability to return high assigned accuracies using less genetic information in a relatively fast period of time. Taken together, these attributes result in reduced expense and reduced time for both data acquisition and computation.

Table 2.1. Diversity measures of microsatellite loci used in simulations. All simulated scenarios used only the primary loci set (#1-8) with the exception of the scenario varying the number of loci used, which conducted separate analyses using the first 4, 8, 12, and 16 loci as numbered. H_e = Expected heterozygosity.

Primary Loci				Supplemental Loci			
#	**Locus**	**Alleles**	**H_e**	**#**	**Locus**	**Alleles**	**H_e**
1	C113	11	0.80	9	Locus 9	10	0.82
2	D75	13	0.79	10	Locus 10	7	0.80
3	C88	8	0.76	11	Locus 11	12	0.79
4	D100	12	0.85	12	Locus 12	8	0.80
5	C115	21	0.86	13	Locus 13	13	0.85
6	C129	5	0.65	14	Locus 14	15	0.87
7	C24	6	0.72	15	Locus 15	7	0.76
8	D237	24	0.87	16	Locus 16	9	0.79

Table 2.2. Parentage assignment classifications and their use in accuracy assessment measures.

True parent sampled	Parent Assigned	Assignment Correct	Symbol
Yes	Yes	Yes	*a*
		No	*b*
	No	No	*c*
No	Yes	No	*d*
	No	Yes	*e*

Accuracy Measure	Equation
Total (TA)	$\frac{a+e}{a+b+c+d+e}$
Assigned (AA)	$\frac{a}{a+b+d}$
True Parent Sampled Correctly Assigned (SA)	$\frac{a}{a+b+c}$
True Parent Not Sampled Incorrectly Assigned (NI)	$\frac{d}{a+b+d}$

Table 2.3. Means and 95% confidence intervals for percent differences (SC-PA/SC) in parentage assignment accuracies between the sibship constraint (SC) method and the raw parentage assignment (PA) output averaged over all eight simulation and analysis variables investigated.

Program	TA	AA	SA	NI	90% CERVUS AA	90% CERVUS NI
Pedapp	11.9 (4.1,19.8)	39.3 (30.4,48.3)	-8.9 (-12.1,-5.7)	-80.3 (-86.8,-73.8)	8.1 (5.9,10.3)	352.7 (217.5,487.8)
Pasos	22.3 (12.6,32.0)	53.1 (42.8,63.3)	-9.0 (-12.4,-5.6)	-82.5 (-88.8,-76.1)	7.0 (4.3,9.7)	346.4 (209.9,482.8)
Cervus	15.2 (12.2,18.2)	49.9 (44.7,55.1)	-20.6 (-26.1,-15.2)	-84.5 (-88.3,-80.6)	7.3 (4.5,10.1)	124.9 (104.1,145.7)

Table 2.4 Effect of genetic sibship reconstruction (SR) program and minimum threshold value on sibship constrained (SC) and raw parentage assignment (PA) accuracies. Values represent mean and 95% confidence intervals of ten replicates. 'Cervus' under the 'Parentage Assignment Program' column refers to output produced using a relaxed 50% confidence level. TA = Total Accuracy, AA = Assigned Accuracy, SA = True Parent Sampled and Assigned Correctly, NI = True Parent Not Sampled and Assigned Incorrectly.

		Sibship Constrained Parentage Assignments			
Sibship Program	**Parentage Assignment Program**	**TA**	**AA**	**SA**	**NI**
Colony	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
Kingroup	Pedapp	0.75 (0.72,0.77)	0.87 (0.84,0.89)	0.61 (0.56,0.66)	0.05 (0.03,0.07)
	Pasos	0.73 (0.70,0.77)	0.86 (0.84,0.89)	0.59 (0.53,0.65)	0.05 (0.03,0.07)
	Cervus	0.59 (0.56,0.63)	0.85 (0.80,0.89)	0.35 (0.26,0.43)	0.08 (0.03,0.12)
Pedigree	Pedapp	0.79 (0.75,0.83)	0.92 (0.89,0.95)	0.67 (0.61,0.73)	0.03 (0.01,0.04)
	Pasos	0.78 (0.74,0.82)	0.92 (0.89,0.95)	0.66 (0.60,0.72)	0.03 (0.01,0.05)
	Cervus	0.62 (0.58,0.66)	0.91 (0.88,0.95)	0.38 (0.30,0.47)	0.06 (0.03,0.08)
Threshold Value					
0	Pedapp	0.81 (0.77,0.86)	0.74 (0.69,0.80)	0.88 (0.84,0.92)	0.16 (0.13,0.20)
	Pasos	0.76 (0.73,0.79)	0.69 (0.65,0.73)	0.87 (0.85,0.90)	0.22 (0.19,0.25)
	Cervus	0.71 (0.67,0.75)	0.69 (0.63,0.75)	0.68 (0.62,0.75)	0.17 (0.13,0.22)
0.1667	Pedapp	0.86 (0.82,0.91)	0.84 (0.78,0.89)	0.87 (0.82,0.91)	0.09 (0.06,0.13)
	Pasos	0.83 (0.80,0.86)	0.80 (0.76,0.83)	0.85 (0.81,0.88)	0.12 (0.09,0.15)
	Cervus	0.70 (0.66,0.75)	0.81 (0.74,0.88)	0.59 (0.51,0.66)	0.13 (0.07,0.18)
0.2501	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
0.3334	Pedapp	0.81 (0.76,0.85)	0.98 (0.96,0.99)	0.68 (0.60,0.76)	0.01 (0.00,0.02)
	Pasos	0.78 (0.74,0.83)	0.97 (0.95,0.99)	0.64 (0.57,0.72)	0.01 (0.00,0.02)
	Cervus	0.60 (0.56,0.64)	0.95 (0.91,0.98)	0.33 (0.24,0.43)	0.04 (0.01,0.08)

Table 2.4 (continued)

Sibship Program	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
Colony	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
Kingroup	Pedapp	0.78 (0.74,0.82)	0.69 (0.63,0.75)	0.83 (0.79,0.87)	0.17 (0.13,0.20)
	Pasos	0.69 (0.66,0.73)	0.62 (0.57,0.67)	0.81 (0.78,0.84)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.60)	0.62 (0.56,0.68)	0.55 (0.48,0.62)	0.33 (0.27,0.38)
Pedigree	Pedapp	0.78 (0.74,0.82)	0.69 (0.64,0.75)	0.83 (0.79,0.87)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.73)	0.62 (0.57,0.67)	0.81 (0.78,0.84)	0.24 (0.20,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.68)	0.55 (0.48,0.63)	0.33 (0.27,0.38)
Threshold Value					
0	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.1667	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.2501	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.3334	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)

Table 2.4 (continued)

		90% Cervus Assignments	
Sibship Program	**Parentage Assignment Program**	**AA**	**SA**
Colony	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
Kingroup	Pedapp		
	Pasos		
	Cervus	0.85 (0.79,0.92)	0.20 (0.14,0.27)
Pedigree	Pedapp		
	Pasos		
	Cervus	0.85 (0.77,0.93)	0.20 (0.13,0.26)
Threshold Value			
0	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.1667	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.2501	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.3334	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)

Table 2.5. Effect of mating strategy and mean sampled full-sibling family size on sibship constrained (SC) and raw parentage assignment (PA) accuracies. Values represent mean and 95% confidence intervals of ten replicates. 'Cervus' under the 'Parentage Assignment Program' column refers to output produced using a relaxed 50% confidence level. TA = Total Accuracy, AA = Assigned Accuracy, SA = True Parent Sampled and Assigned Correctly, NI = True Parent Not Sampled and Assigned Incorrectly.

		Sibship Constrained Parentage Assignments			
Mating Strategy	**Parentage Assignment Program**	**TA**	**AA**	**SA**	**NI**
Monogamy	Pedapp	0.88 (0.86,0.91)	0.94 (0.91,0.97)	0.83 (0.80,0.86)	0.03 (0.01,0.06)
	Pasos	0.86 (0.83,0.88)	0.92 (0.89,0.95)	0.82 (0.77,0.86)	0.06 (0.03,0.09)
	Cervus	0.67 (0.64,0.71)	0.93 (0.91,0.96)	0.47 (0.40,0.55)	0.05 (0.03,0.06)
Polyandry	Pedapp	0.87 (0.86,0.89)	0.94 (0.92,0.96)	0.80 (0.77,0.83)	0.03 (0.01,0.05)
	Pasos	0.84 (0.82,0.86)	0.94 (0.91,0.97)	0.75 (0.72,0.79)	0.03 (0.02,0.05)
	Cervus	0.68 (0.64,0.71)	0.94 (0.91,0.97)	0.46 (0.40,0.52)	0.05 (0.02,0.08)
Polygamy	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
Mean Family Size					
1.9	Pedapp	0.80 (0.77,0.84)	0.94 (0.92,0.97)	0.68 (0.64,0.73)	0.03 (0.01,0.04)
	Pasos	0.79 (0.75,0.82)	0.92 (0.89,0.95)	0.66 (0.62,0.71)	0.04 (0.02,0.07)
	Cervus	0.65 (0.62,0.67)	0.93 (0.90,0.96)	0.41 (0.36,0.45)	0.04 (0.02,0.07)
2.5	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
3.5	Pedapp	0.89 (0.87,0.92)	0.96 (0.92,0.99)	0.83 (0.77,0.88)	0.02 (0.01,0.04)
	Pasos	0.87 (0.84,0.89)	0.95 (0.93,0.97)	0.79 (0.74,0.83)	0.03 (0.01,0.05)
	Cervus	0.67 (0.63,0.71)	0.94 (0.89,1.00)	0.42 (0.33,0.51)	0.05 (-0.01,0.10)
4.3	Pedapp	0.91 (0.88,0.95)	0.98 (0.96,1.00)	0.86 (0.78,0.93)	0.01 (-0.01,0.03)
	Pasos	0.90 (0.88,0.92)	0.97 (0.95,0.99)	0.86 (0.81,0.90)	0.02 (0.00,0.04)
	Cervus	0.66 (0.62,0.70)	0.97 (0.94,1.00)	0.43 (0.34,0.52)	0.02 (0.00,0.04)

Table 2.5 (continued)

Mating Strategy	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
Monogamy	Pedapp	0.76 (0.73,0.79)	0.67 (0.62,0.72)	0.84 (0.81,0.87)	0.20 (0.16,0.24)
	Pasos	0.69 (0.64,0.73)	0.61 (0.55,0.67)	0.83 (0.79,0.86)	0.27 (0.22,0.32)
	Cervus	0.57 (0.54,0.61)	0.62 (0.56,0.69)	0.56 (0.51,0.62)	0.33 (0.26,0.39)
Polyandry	Pedapp	0.79 (0.77,0.81)	0.70 (0.67,0.73)	0.86 (0.83,0.89)	0.19 (0.16,0.22)
	Pasos	0.69 (0.66,0.71)	0.60 (0.57,0.64)	0.82 (0.79,0.85)	0.27 (0.24,0.30)
	Cervus	0.57 (0.55,0.59)	0.60 (0.56,0.63)	0.58 (0.54,0.63)	0.35 (0.31,0.38)
Polygamy	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
Mean Family Size					
1.9	Pedapp	0.79 (0.76,0.81)	0.70 (0.67,0.73)	0.86 (0.84,0.87)	0.19 (0.16,0.21)
	Pasos	0.69 (0.67,0.72)	0.62 (0.59,0.64)	0.84 (0.82,0.86)	0.27 (0.25,0.29)
	Cervus	0.56 (0.54,0.59)	0.60 (0.56,0.64)	0.59 (0.55,0.63)	0.35 (0.31,0.38)
2.5	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
3.5	Pedapp	0.77 (0.74,0.81)	0.67 (0.61,0.72)	0.82 (0.79,0.85)	0.19 (0.14,0.23)
	Pasos	0.67 (0.62,0.71)	0.57 (0.51,0.64)	0.79 (0.76,0.82)	0.28 (0.23,0.34)
	Cervus	0.55 (0.52,0.58)	0.57 (0.49,0.65)	0.52 (0.46,0.57)	0.36 (0.29,0.42)
4.3	Pedapp	0.76 (0.73,0.80)	0.67 (0.60,0.74)	0.83 (0.79,0.87)	0.19 (0.13,0.25)
	Pasos	0.68 (0.63,0.72)	0.60 (0.53,0.67)	0.81 (0.77,0.85)	0.27 (0.20,0.34)
	Cervus	0.55 (0.52,0.58)	0.61 (0.52,0.70)	0.53 (0.47,0.58)	0.34 (0.25,0.43)

Table 2.5 (continued)

		90% Cervus Assignments	
Mating Strategy	**Parentage Assignment Program**	**AA**	**SA**
Monogamy	Pedapp		
	Pasos		
	Cervus	0.88 (0.82,0.93)	0.18 (0.13,0.22)
Polyandry	Pedapp		
	Pasos		
	Cervus	0.88 (0.82,0.93)	0.20 (0.14,0.25)
Polygamy	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
Mean Family Size			
1.9	Pedapp		
	Pasos		
	Cervus	0.91 (0.87,0.94)	0.21 (0.17,0.25)
2.5	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
3.5	Pedapp		
	Pasos		
	Cervus	0.86 (0.80,0.91)	0.17 (0.11,0.22)
4.3	Pedapp		
	Pasos		
	Cervus	0.84 (0.74,0.93)	0.15 (0.11,0.19)

Table 2.6. Effect of proportion of actual parents sampled and number of loci used in analyses on sibship constrained (SC) and raw parentage assignment (PA) accuracies. Values represent mean and 95% confidence intervals of ten replicates. 'Cervus' under the 'Parentage Assignment Program' column refers to output produced using a relaxed 50% confidence level. TA = Total Accuracy, AA = Assigned Accuracy, SA = True Parent Sampled and Assigned Correctly, NI = True Parent Not Sampled and Assigned Incorrectly.

		Sibship Constrained Parentage Assignments			
Proportion of Parents Sampled	**Parentage Assignment Program**	**TA**	**AA**	**SA**	**NI**
0.2	Pedapp	0.93 (0.91,0.94)	0.82 (0.75,0.89)	0.82 (0.77,0.86)	0.16 (0.10,0.23)
	Pasos	0.92 (0.91,0.93)	0.80 (0.74,0.85)	0.82 (0.77,0.86)	0.19 (0.13,0.25)
	Cervus	0.83 (0.79,0.86)	1.00 (1.00,1.00)	0.15 (0.06,0.25)	0.00 (0.00,0.00)
0.4	Pedapp	0.88 (0.87,0.90)	0.86 (0.81,0.91)	0.77 (0.72,0.82)	0.10 (0.05,0.15)
	Pasos	0.89 (0.86,0.91)	0.89 (0.85,0.92)	0.78 (0.74,0.82)	0.09 (0.05,0.13)
	Cervus	0.70 (0.64,0.76)	0.94 (0.88,1.00)	0.20 (0.12,0.28)	0.03 (-0.01,0.07)
0.6	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
0.8	Pedapp	0.84 (0.82,0.87)	0.95 (0.92,0.97)	0.81 (0.77,0.85)	0.02 (0.00,0.03)
	Pasos	0.83 (0.80,0.85)	0.95 (0.92,0.97)	0.79 (0.74,0.84)	0.02 (0.00,0.04)
	Cervus	0.81 (0.79,0.84)	0.93 (0.91,0.95)	0.79 (0.74,0.83)	0.04 (0.02,0.06)
0.98	Pedapp	0.85 (0.82,0.88)	0.98 (0.97,0.99)	0.84 (0.81,0.88)	0.00 (0.00,0.00)
	Pasos	0.83 (0.79,0.86)	0.98 (0.97,0.99)	0.82 (0.79,0.86)	0.00 (0.00,0.00)
	Cervus	0.83 (0.79,0.87)	0.98 (0.96,0.99)	0.83 (0.78,0.87)	0.00 (0.00,0.00)

Table 2.6. (continued)

Proportion of Parents Sampled	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
0.2	Pedapp	0.77 (0.76,0.79)	0.44 (0.39,0.49)	0.88 (0.86,0.90)	0.50 (0.44,0.55)
	Pasos	0.71 (0.69,0.73)	0.38 (0.34,0.42)	0.86 (0.83,0.88)	0.57 (0.52,0.61)
	Cervus	0.82 (0.79,0.85)	0.62 (0.50,0.74)	0.22 (0.14,0.29)	0.36 (0.25,0.47)
0.4	Pedapp	0.75 (0.72,0.79)	0.55 (0.47,0.63)	0.84 (0.82,0.86)	0.34 (0.25,0.44)
	Pasos	0.66 (0.63,0.69)	0.48 (0.41,0.55)	0.84 (0.81,0.87)	0.44 (0.35,0.52)
	Cervus	0.65 (0.60,0.70)	0.52 (0.44,0.60)	0.36 (0.30,0.42)	0.45 (0.37,0.53)
0.6	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.8	Pedapp	0.81 (0.79,0.84)	0.77 (0.74,0.81)	0.85 (0.83,0.87)	0.09 (0.06,0.12)
	Pasos	0.74 (0.70,0.78)	0.71 (0.66,0.76)	0.83 (0.80,0.87)	0.15 (0.11,0.19)
	Cervus	0.64 (0.60,0.69)	0.65 (0.60,0.69)	0.83 (0.81,0.86)	0.23 (0.19,0.28)
0.98	Pedapp	0.87 (0.86,0.89)	0.87 (0.86,0.89)	0.88 (0.87,0.89)	0.01 (0.00,0.02)
	Pasos	0.85 (0.83,0.86)	0.85 (0.83,0.86)	0.86 (0.85,0.87)	0.02 (0.01,0.02)
	Cervus	0.84 (0.82,0.86)	0.85 (0.83,0.87)	0.86 (0.84,0.87)	0.02 (0.01,0.03)

Table 2.6. (continued)

		90% Cervus Assignments	
Proportion of Parents Sampled	**Parentage Assignment Program**	**AA**	**SA**
0.2	Pedapp		
	Pasos		
	Cervus	0.82 (0.60,1.04)	0.04 (0.00,0.08)
0.4	Pedapp		
	Pasos		
	Cervus	0.78 (0.66,0.89)	0.06 (0.05,0.08)
0.6	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.8	Pedapp		
	Pasos		
	Cervus	0.90 (0.87,0.93)	0.40 (0.35,0.44)
0.98	Pedapp		
	Pasos		
	Cervus	0.92 (0.91,0.94)	0.70 (0.66,0.74)

Table 2.6. (continued)

Number of Loci Used	Parentage Assignment Program	Sibship Constrained Parentage Assignments			
		TA	AA	SA	NI
4	Pedapp	0.49 (0.45,0.53)	0.48 (0.43,0.52)	0.23 (0.20,0.25)	0.17 (0.12,0.22)
	Pasos	0.49 (0.44,0.53)	0.48 (0.41,0.54)	0.20 (0.18,0.23)	0.16 (0.11,0.21)
	Cervus	0.41 (0.36,0.45)	N/A	N/A	N/A
8	Pedapp	0.86 (0.85,0.88)	0.95 (0.93,0.97)	0.79 (0.77,0.81)	0.03 (0.02,0.04)
	Pasos	0.84 (0.82,0.86)	0.93 (0.91,0.95)	0.76 (0.74,0.79)	0.05 (0.03,0.07)
	Cervus	0.66 (0.62,0.69)	0.96 (0.94,0.98)	0.43 (0.40,0.47)	0.04 (0.01,0.06)
12	Pedapp	0.96 (0.95,0.97)	0.98 (0.96,1.00)	0.94 (0.93,0.95)	0.01 (0.00,0.02)
	Pasos	0.96 (0.95,0.97)	0.98 (0.97,1.00)	0.94 (0.93,0.95)	0.02 (0.00,0.03)
	Cervus	0.86 (0.83,0.89)	0.93 (0.91,0.95)	0.82 (0.76,0.87)	0.06 (0.04,0.09)
16	Pedapp	0.99 (0.98,1.00)	0.99 (0.99,1.00)	0.98 (0.97,0.99)	0.00 (0.00,0.01)
	Pasos	0.99 (0.98,0.99)	0.99 (0.99,1.00)	0.98 (0.97,0.99)	0.00 (0.00,0.01)
	Cervus	0.90 (0.88,0.92)	0.90 (0.88,0.92)	0.92 (0.89,0.95)	0.09 (0.07,0.12)

Table 2.6. (continued)

Number of Loci Used	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
4	Pedapp	0.22 (0.21,0.24)	0.20 (0.19,0.22)	0.34 (0.31,0.36)	0.39 (0.35,0.44)
	Pasos	0.19 (0.18,0.21)	0.19 (0.17,0.20)	0.31 (0.29,0.34)	0.40 (0.36,0.45)
	Cervus	0.41 (0.36,0.46)	0.43 (0.08,0.77)	0.01 (0.00,0.02)	0.23 (0.05,0.4)
8	Pedapp	0.77 (0.75,0.78)	0.68 (0.66,0.71)	0.85 (0.82,0.87)	0.20 (0.17,0.22)
	Pasos	0.69 (0.67,0.71)	0.61 (0.59,0.64)	0.83 (0.80,0.85)	0.26 (0.24,0.29)
	Cervus	0.57 (0.55,0.58)	0.62 (0.58,0.67)	0.56 (0.54,0.59)	0.33 (0.29,0.37)
12	Pedapp	0.93 (0.91,0.95)	0.89 (0.86,0.92)	0.96 (0.94,0.97)	0.07 (0.05,0.10)
	Pasos	0.91 (0.89,0.93)	0.87 (0.84,0.89)	0.96 (0.95,0.97)	0.11 (0.08,0.13)
	Cervus	0.66 (0.63,0.69)	0.64 (0.60,0.68)	0.86 (0.83,0.88)	0.33 (0.29,0.37)
16	Pedapp	0.98 (0.97,0.99)	0.97 (0.95,0.99)	0.99 (0.99,1.00)	0.02 (0.01,0.04)
	Pasos	0.97 (0.96,0.98)	0.96 (0.95,0.97)	0.99 (0.98,0.99)	0.03 (0.02,0.04)
	Cervus	0.73 (0.71,0.75)	0.70 (0.67,0.73)	0.92 (0.90,0.95)	0.29 (0.26,0.32)

Table 2.6. (continued)

		90% Cervus Assignments	
Number of Loci Used	**Parentage Assignment Program**	**AA**	**SA**
4	Pedapp		
	Pasos		
	Cervus	N/A	N/A
8	Pedapp		
	Pasos		
	Cervus	0.89 (0.83,0.96)	0.20 (0.16,0.24)
12	Pedapp		
	Pasos		
	Cervus	0.90 (0.88,0.91)	0.55 (0.49,0.61)
16	Pedapp		
	Pasos		
	Cervus	0.88 (0.86,0.91)	0.67 (0.62,0.72)

Table 2.7. Effect of genetic and demographic error rates on sibship constrained (SC) and raw parentage assignment (PA) accuracies. Values represent mean and 95% confidence intervals of ten replicates. 'Cervus' under the 'Parentage Assignment Program' column refers to output produced using a relaxed 50% confidence level. TA = Total Accuracy, AA = Assigned Accuracy, SA = True Parent Sampled and Assigned Correctly, NI = True Parent Not Sampled and Assigned Incorrectly

		Sibship Constrained Parentage Assignments			
Genetic Error	**Parentage Assignment Program**	**TA**	**AA**	**SA**	**NI**
0	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
0.01	Pedapp	0.83 (0.79,0.86)	0.95 (0.93,0.96)	0.71 (0.65,0.77)	0.02 (0.00,0.03)
	Pasos	0.79 (0.76,0.83)	0.93 (0.90,0.96)	0.66 (0.59,0.73)	0.03 (0.01,0.05)
	Cervus	0.66 (0.61,0.70)	0.95 (0.91,0.98)	0.41 (0.34,0.48)	0.03 (0.01,0.06)
0.03	Pedapp	0.76 (0.72,0.79)	0.93 (0.90,0.96)	0.60 (0.55,0.66)	0.04 (0.01,0.07)
	Pasos	0.78 (0.75,0.80)	0.93 (0.91,0.95)	0.63 (0.57,0.68)	0.03 (0.01,0.05)
	Cervus	0.62 (0.58,0.66)	0.91 (0.87,0.95)	0.35 (0.30,0.41)	0.07 (0.02,0.13)
0.05	Pedapp	0.71 (0.67,0.76)	0.94 (0.91,0.97)	0.53 (0.46,0.60)	0.03 (0.01,0.05)
	Pasos	0.72 (0.68,0.75)	0.92 (0.88,0.96)	0.54 (0.49,0.60)	0.04 (0.02,0.06)
	Cervus	0.52 (0.48,0.56)	0.91 (0.85,0.97)	0.20 (0.14,0.26)	0.05 (0.02,0.08)

Table 2.7. (continued)

Genetic Error	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
0	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.01	Pedapp	0.76 (0.74,0.79)	0.67 (0.62,0.72)	0.79 (0.76,0.82)	0.17 (0.13,0.22)
	Pasos	0.68 (0.64,0.72)	0.59 (0.53,0.64)	0.77 (0.74,0.80)	0.26 (0.20,0.31)
	Cervus	0.54 (0.51,0.56)	0.56 (0.49,0.63)	0.55 (0.52,0.58)	0.37 (0.30,0.43)
0.03	Pedapp	0.71 (0.69,0.73)	0.64 (0.59,0.68)	0.70 (0.66,0.73)	0.19 (0.16,0.21)
	Pasos	0.65 (0.63,0.67)	0.58 (0.54,0.62)	0.72 (0.69,0.75)	0.26 (0.22,0.29)
	Cervus	0.56 (0.52,0.59)	0.61 (0.55,0.67)	0.49 (0.45,0.53)	0.32 (0.27,0.38)
0.05	Pedapp	0.67 (0.64,0.70)	0.61 (0.57,0.65)	0.63 (0.58,0.67)	0.18 (0.15,0.20)
	Pasos	0.62 (0.59,0.66)	0.56 (0.51,0.60)	0.65 (0.60,0.69)	0.25 (0.22,0.27)
	Cervus	0.51 (0.49,0.54)	0.60 (0.57,0.63)	0.38 (0.33,0.42)	0.32 (0.29,0.35)

Table 2.7. (continued)

		90% Cervus Assignments	
Genetic Error	**Parentage Assignment Program**	**AA**	**SA**
0	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.01	Pedapp		
	Pasos		
	Cervus	0.85 (0.80,0.90)	0.19 (0.16,0.22)
0.03	Pedapp		
	Pasos		
	Cervus	0.89 (0.82,0.95)	0.15 (0.13,0.17)
0.05	Pedapp		
	Pasos		
	Cervus	0.88 (0.81,0.95)	0.08 (0.06,0.11)

Table 2.7. (continued)

		Sibship Constrained Parentage Assignments			
Demographic Error	**Parentage Assignment Program**	**TA**	**AA**	**SA**	**NI**
0	Pedapp	0.86 (0.83,0.88)	0.95 (0.92,0.98)	0.77 (0.71,0.82)	0.02 (0.00,0.03)
	Pasos	0.84 (0.81,0.87)	0.95 (0.92,0.97)	0.75 (0.69,0.81)	0.02 (0.01,0.03)
	Cervus	0.66 (0.62,0.71)	0.94 (0.90,0.98)	0.45 (0.35,0.54)	0.04 (0.01,0.08)
0.05	Pedapp	0.85 (0.83,0.87)	0.95 (0.92,0.98)	0.76 (0.72,0.80)	0.03 (0.00,0.06)
	Pasos	0.84 (0.81,0.86)	0.92 (0.88,0.96)	0.76 (0.71,0.80)	0.05 (0.01,0.10)
	Cervus	0.66 (0.61,0.70)	0.93 (0.88,0.98)	0.44 (0.37,0.50)	0.04 (0.00,0.08)
0.1	Pedapp	0.81 (0.78,0.85)	0.93 (0.89,0.96)	0.73 (0.67,0.78)	0.05 (0.02,0.08)
	Pasos	0.79 (0.76,0.82)	0.91 (0.89,0.93)	0.70 (0.64,0.75)	0.06 (0.03,0.08)
	Cervus	0.65 (0.61,0.69)	0.94 (0.91,0.96)	0.44 (0.37,0.51)	0.04 (0.02,0.07)
0.15	Pedapp	0.79 (0.76,0.83)	0.94 (0.92,0.96)	0.68 (0.61,0.74)	0.03 (0.01,0.04)
	Pasos	0.79 (0.76,0.82)	0.92 (0.90,0.95)	0.69 (0.62,0.76)	0.04 (0.03,0.06)
	Cervus	0.61 (0.57,0.65)	0.91 (0.88,0.94)	0.38 (0.29,0.47)	0.06 (0.04,0.09)

Table 2.7. (continued)

Demographic Error	Parentage Assignment Program	Raw Parentage Assignments			
		TA	AA	SA	NI
0	Pedapp	0.78 (0.74,0.83)	0.69 (0.63,0.76)	0.83 (0.78,0.88)	0.17 (0.13,0.20)
	Pasos	0.70 (0.66,0.74)	0.62 (0.57,0.68)	0.81 (0.77,0.85)	0.24 (0.21,0.28)
	Cervus	0.56 (0.52,0.61)	0.61 (0.55,0.67)	0.56 (0.48,0.63)	0.33 (0.28,0.39)
0.05	Pedapp	0.78 (0.75,0.81)	0.69 (0.64,0.75)	0.85 (0.83,0.87)	0.18 (0.13,0.24)
	Pasos	0.68 (0.64,0.72)	0.61 (0.56,0.67)	0.83 (0.80,0.86)	0.27 (0.21,0.33)
	Cervus	0.58 (0.55,0.60)	0.62 (0.58,0.67)	0.57 (0.53,0.61)	0.31 (0.27,0.36)
0.1	Pedapp	0.76 (0.72,0.79)	0.67 (0.61,0.74)	0.84 (0.80,0.88)	0.20 (0.14,0.26)
	Pasos	0.67 (0.62,0.72)	0.60 (0.54,0.67)	0.81 (0.77,0.85)	0.27 (0.20,0.33)
	Cervus	0.58 (0.54,0.61)	0.63 (0.55,0.71)	0.60 (0.54,0.66)	0.32 (0.25,0.39)
0.15	Pedapp	0.76 (0.72,0.80)	0.68 (0.62,0.74)	0.82 (0.78,0.85)	0.17 (0.13,0.21)
	Pasos	0.67 (0.62,0.71)	0.60 (0.54,0.65)	0.80 (0.76,0.83)	0.26 (0.21,0.30)
	Cervus	0.55 (0.53,0.58)	0.62 (0.56,0.68)	0.55 (0.51,0.6)	0.32 (0.26,0.37)

Table 2.7. (continued)

		90% Cervus Assignments	
Demographic Error	**Parentage Assignment Program**	**AA**	**SA**
0	Pedapp		
	Pasos		
	Cervus	0.84 (0.76,0.92)	0.21 (0.14,0.28)
0.05	Pedapp		
	Pasos		
	Cervus	0.90 (0.86,0.94)	0.18 (0.15,0.21)
0.1	Pedapp		
	Pasos		
	Cervus	0.90 (0.84,0.96)	0.20 (0.17,0.23)
0.15	Pedapp		
	Pasos		
	Cervus	0.86 (0.81,0.91)	0.18 (0.13,0.22)

CHAPTER III

MATING PATTERNS IN WILD BROOK TROUT POPULATIONS: INFLUENCES ON EFFECTIVE POPULATION SIZE AND IMPLICATIONS FOR POPULATION PERSISTENCE

Abstract

Sexual selection directly influences effective population size (N_e) by determining variance in reproductive success. The proportion of mature adults and the frequency with which they successfully reproduce determines the rate at which genetic variation is lost from a population. In salmonids, there is intense competition amongst males to acquire mating privileges from females, with total body size typically indicative of competitive superiority. Furthermore, female spawning is temporally asynchronous progressing from larger to smaller sized individuals. Given these two factors, opportunity exists for large dominant males to monopolize mature females over their entire size range. Alternatively, mate choice may be size-assortative with either females selecting males of similar size, or males only competing for females of similar size. These two scenarios have vastly different implications for the resulting N_e. Using genetically reconstructed pedigrees for two wild brook trout (*Salvelinus fontinalis*) populations, we measured correlation in size of mating pairs, total number of individuals contributing to reproduction, and individual success in terms of number of full-sibling families and offspring produced. Mating pairs were significantly size-assortative, with individual length accounting for 37% of the

variation. This pattern of size-assortative mate choice resulted in a reproductive strategy closer to monogamy than polygamy. Of all reproducing adults (n=157), 80% (n=126) produced only one full-sibling family, and only 6% (n= 9) contributed to two or more full-sibling families. The number of families and offspring contributed increased with length for both males and females. Comparison of the effective population size estimate to the adult census size estimate returned an N_e/N_c ratio of 0.49 averaged over both populations. This value is nearly five times greater than the average reported across 165 (0.14) and 102 (0.10) different species. These findings are discussed in the context of population persistence given the trend of increasing habitat fragmentation, and expected change in hydrologic regimes caused by climate change.

Introduction

The level of genetic variation within a population influences its fitness and viability, and ultimately determines its evolutionary potential. Reduced genetic variation has been empirically demonstrated to directly decrease fitness in both laboratory (Frankham 2005) and wild (Crnokrak & Roff 1999;Keller & Waller 2002) populations, and negatively affect population viability (Newman & Pilson 1997;Reed 2005;Saccheri *et al*. 1998). Additionally, reduced genetic diversity has been shown to limit the adaptive potential of populations subjected to altered environmental conditions (Frankham *et al*. 1999).

Genetic variation is lost when the assumptions of an 'ideal' population (random mating, constant size, equal sex ratio, discrete generations) are violated (Wright 1931).

These violations result in the population behaving as one of a smaller size in terms of the rate of change in allele frequencies or heterozygosity (Luikart *et al*. 2010). This smaller size is defined as the effective population size (N_e), and is one of the fundamental metrics in both evolutionary and conservation biology (Charlesworth 2009;Frankham 2005).

Violation of random mating, which results in increased reproductive variance, has been attributed as one of the primary causes of reduced N_e both theoretically (Hedrick 2005) and empirically (Araki *et al*. 2007;Frankham 1995). Non-random mating arises when members of either sex select mates based on specific traits (Clutton-Brock 2007), and choice patterns can be either dominant (Bateman 1948) or assortative (Crespi 1989) relative to that trait. These two patterns result in substantially different reproductive variances, with dominant patterns resulting in lower N_e caused by greater skew in parental contribution (Wade & Arnold 1980).

In salmonids, opportunity exists for dominant patterns to occur. Males compete for mating rights with females, with size typically indicating superiority (Blanchfield & Ridgway 1999;Blanchfield *et al.* 2003;Fleming 1996;Jones & Hutchings 2001;Labonne *et al*. 2009). Furthermore, female spawning is temporally asynchronous, progressing from larger to smaller sized individuals (Blanchfield & Ridgway 2005;Elliott 1984). Given these two factors, opportunity exists for large dominant males to monopolize mature females over their entire size range.

Alternatively, mate choice may be size-assortative. This type of pattern could arise if females prefer larger males, and males only compete for females of equal or larger size. This would limit potential pairs to the overlap between the two groups, resulting in mates of approximately equal size. These sex-specific behavioral preferences

have been observed in wild salmonid populations (Blanchfield & Ridgway 1999;Esteve 2005;Labonne *et al.* 2009).

Empirical evidence for the type of mating strategy employed by salmonids in wild populations is equivocal. Blanchfield and Ridgway (1999) reported size-assortative mating in a lacustrine population of brook trout, and Taggart *et al*. (2001) reported the same pattern for one cohort of Atlantic salmon. In contrast, Dickerson *et al*. (2004) found evidence supporting a dominant mating pattern in a population of pink salmon. Thus, more research is required in this area before conclusions can be drawn.

The importance of understanding mating strategies becomes more apparent as the number of salmonid populations affected by anthropogenic impacts rises, and informed conservation and management decisions are required. Possibly of greatest concern are resident salmonid populations inhabiting headwater stream environments. Historically they have been impacted by land use practices (Nislow & Lowe 2003), reduced water quality (Hudy *et al.* 2000), and increased habitat fragmentation (Wofford *et al.* 2005). Hudy *et al*. (2008) examined the distribution and status of wild brook trout populations over their native range in the U.S. and reported that 35% of sub-watersheds contained less than half of their historic habitat, and 28% suffered population extirpation. Additionally, future hydrologic conditions within watersheds are expected to be negatively impacted by climate change (Marshall & Randhir 2008). These data highlight the need for brook trout conservation strategies, an important component of which is maintenance of genetic variation.

As habitat fragmentation increases and population sizes decrease, it becomes imperative to understand population dynamics in headwater systems in order to predict

loss of genetic variation and estimate vulnerability. Given this, the objectives for this study were to 1) determine mating strategies employed by brook trout populations inhabiting headwater stream environments, 2) calculate the effective population size and its ratio to census size, and 3) relate results to conservation and management strategies for population persistence.

Materials and Methods

Study Sites and Sampling Design

The first brook trout population was located in the Fridley Gap (FG) watershed in Rockingham County, VA, USA (Figure 3.1A). Watershed area was 5.6 km2. Specific geology and land use history are summarized in Hudy *et al*. (2000). The Fridley Gap study area was a 1.8 km stretch of stream with one tributary (250 M long) entering the main stem at river km 1.5. The downstream end of the study area was bounded by a small, impassable dam, while the upstream ends of both the main stem and tributary were limited by intermittent flows. The habitat below the dam was impacted by agricultural practices and suffered from severely degraded riparian habitat making it unsuitable for brook trout. The average low-flow wetted width for the main stem was 3.8 m, while the average low-flow wetted width for the tributary was 1.8 m.

The second brook trout population was located in the West Brook (WB) watershed in Whately, MA, USA (Figure 3.1B). Watershed area was 11.8 km2. Stream habitat and land use are described in Letcher *et al*. (2002; 2007). The West Brook study

area was an approximately 1 km stretch of stream with two tributaries entering into the main stem at river kilometers 0.4 (Open-Small (OS)) and 0.6 (Open-Large (OL)). The downstream end of the study area was bounded by a small waterfall (1-m tall, passable by fish), while the upstream end was unbounded. Each tributary study area was 300 m long and were both bounded by impassable waterfalls at their upstream ends. The confluence of the OL tributary was open, but the confluence of the OS tributary was interrupted by a perched culvert (≈1 m tall, passable by fish (Letcher *et al*. 2007). Average wetted widths of the three streams were 4.5 m (WB), 2 m (OS), and 3 m (OL).

Both brook trout populations were maintained through natural reproduction during the study period. However, the FG population was supplemented with 91 individuals taken from a nearby wild population in 1993 after habitat mitigation was performed on the watershed (Hudy *et al.* 2000). The WB population was historically stocked with hatchery reared individuals, however annual stocking ceased in 1997, and there is no evidence for hatchery introgression. Fishing pressure was very low for both populations.

Sampling of the FG population consisted of using single-pass electrofishing for the entire study area during July 2004, 2005, and 2006. We recorded length (± 1 mm fork length), location (nearest m), and collected anal fin clips for genetic analysis from all brook trout captured. We sampled in July because at that time young-of-year (YOY) brook trout were large enough to be efficiently captured by electrofishing, but still small enough to differentiate from older age classes based on length. Estimated capture probability for adult brook trout (over-yearlings) based on capture-mark-recapture experiments was 0.65 (M. Hudy, unpublished data).

Sampling of the WB population was conducted from June 2002 to December 2006. Samples were seasonal, typically in March, June, September and December, although ice buildup precluded December sampling in the main stem four of the five years, and in the tributaries one of the five years. A total of 18 sampling occasions were completed during the study period. We used standard two-pass electrofishing (300V unpulsed DC current) with block nets at the upstream and downstream ends of each 20-m long sampling section in the main-stem, and single-pass electrofishing without block nets in the tributaries. Upon capture, we took lengths (± 1 mm fork length), weights (± 0.1 mg wet weight), and recorded the sampling location (section) and the maturity status and sex (if caught during the fall) for each fish. Untagged fish were implanted with 12-mm PIT tags (Digital Angel, St. Paul MN, USA) if fork length exceeded 60 mm (Gries & Letcher 2002). Anal fin clips for genetic analysis were taken on all untagged fish. Following work-up, fish were returned to their capture location. Capture probability for all seasons and all age classes was 0.6 (Letcher *et al.* 2007).

Genetic Diversity and Genotyping Error

Because brook trout are cryptic breeders with no parental care, we used individual genotypes to reconstruct the pedigree structure among sampled individuals. Panels of twelve (WB) (SfoB52, SfoC24, SfoC38, SfoC86, SfoC88, SfoC113, SfoC115, SfoC129, SfoD75, SfoD91a, SfoD100 (King *et al*. 2003), SsaD237 (King *et al.* 2005)) and eight (FG) (SfoC24, SfoC88, SfoC113, SfoC115, SfoC129, SfoD75, SfoD100, SsaD237) microsatellite loci were selected based on their ability to accurately reconstruct full-

sibling families and assign parents for synthetic data (see below). Protocols for DNA extraction and amplification followed King *et al*. (2005). Loci were electrophoresed on an ABI Prism 3100-Avant genetic analyzer (Applied Biosystems Inc., Foster City, CA), and alleles were scored using GENESCAN v3.7, GENEMAPPER v3.2 and PEAK SCANNER v1.0 software (Applied Biosystems Inc.).

Standard measures of genetic diversity were calculated to assess marker quality. Allele number and observed and expected heterozygosities were calculated using GDA v1.0 (Lewis & Zaykin 2001). Estimation of fis, an analogue of Wright's FIS, and testing for departures from Hardy-Weinberg equilibrium were performed using GENEPOP v4.0.10 (Rousset 2008). Testing was conducted using the heterozygote deficiency option because the presence of a null allele was suspected for at least one locus in the FG population. Tests were performed for each locus in each population (FG: k = 24, WB: k = 48) on each cohort, and significance was assessed using a sequential Bonferroni correction (Holm 1979;Rice 1989) with an α of 0.05. For significant loci, null allele frequencies were estimated using ML-Relate v090408 (Kalinowski *et al.* 2006).

Because genotyping error has the potential to bias pedigree reconstruction (Wilson & Ferguson 2002), genotyping error rates were assessed and used to minimize its impact. For the WB population, 100 individuals were randomly selected to undergo a second DNA extraction and amplification of all twelve loci. Alleles were compared between the two genotypes for each individual and a per allele error rate estimate was obtained. For the FG population, a subset of individuals captured during the 2004 sample were implanted with PIT tags (Hudy *et al*. 2010). This allowed for direct comparison of genotypes for individuals recaptured during a subsequent sampling occasion.

Pedigree Reconstruction

The power of the loci panel to reconstruct full-sibling families and assign parents accurately was assessed through the use of synthetic data generated by the program PEDAGOG v1.2 (Coombs *et al.* 2010a). Demographic parameters for the simulated populations were derived from field data for the WB brook trout population. Genetic parameters (loci number and allele frequencies) were derived using genotyped individuals from the population being simulated. Each simulated population was subjected to a sampling scheme (annual (FG), seasonal (WB)) also using field derived capture probability estimates. Sibship reconstruction and initial parentage assignment analyses were performed on the simulated population using the programs COLONY v1.2 (Wang 2004) (sibship) and PEDAPP v1.1 (Almudevar 2007) (parentage). Final parentage assignments were acquired using the sibship constraint (SC) method within the program PEDAGREE v1.04 (Coombs *et al.* 2010b). For WB populations, the SC method was run using a minimum threshold value of 0.2501 for full-sibling families with two members, and 0.1667 for full-sibling families with three or more members. The results from the two runs were then merged. For FG populations, a minimum threshold value of 0.2501 was used for all full-sibling families with two or more members. Accuracy of reconstructed families and assigned parents were calculated using PEDAGREE. A total of ten replicates were simulated for each population. The same methodology outlined above was used to construct pedigrees for the WB and FG empirical datasets.

Census and Effective Population Size Estimation

Multiple sampling occasions allowed the use of both single-sample and temporal effective population size (N_e) estimators. The linkage disequilibrium (LD) method (Waples 2006) within the program LDNe (Waples & Do 2008) was used to acquire single-sample estimates. Since brook trout have variable age at maturity, the number of cohorts included in the sample was roughly equivalent to the generation length. Waples and Do (2010) conjectured that this estimate should correspond to the N_e for a generation instead of the effective number of breeders (N_b) for a cohort. For the OS and OL populations in the WB drainage, a sample was composed of all brook trout captured during the fall sample, resulting in four samples for each population. For the FG population, samples were composed of all brook trout captured during July of 2004 and 2006. An estimate was not generated for July of 2005 because fin clips were not taken for YOY during that sample. Estimates of N_e were generated for each sample and averaged for each population using a weighted harmonic mean (Waples & Do 2010).

Temporal estimates of N_e were generated using the pseudo maximum-likelihood (ML) method (Wang 2001) within the program MLNe v1.1 (Wang & Whitlock 2003). Samples were delineated into individuals belonging to the same cohort (2002-2005 in WB, 2004 and 2006 in FG) (Jorde & Ryman 1995). For the WB populations, N_e estimates were generated for each consecutive interval (2002-2003, 2003-2004, 2004-2005), and over the entire time period using all cohorts simultaneously. The FG population only contained one interval (2004-2006), and thus produced only one estimate.

For the WB populations, available life-table data (Letcher *et al.* 2007) allowed for estimation of N_e using the generational-overlap correction factor of Jorde and Ryman (1995). N_e estimates were generated using the unbiased F_s estimator (Jorde & Ryman 2007) within the program TempoFS (Available for download at http://www.zoologi.su.se/~ryman/). Samples were collected according to sample plan I (Waples 2005) and N_e estimates were generated using initial population sizes of 62 (OS) and 220 (OL). In addition to an estimate for each consecutive interval, an overall harmonic mean was calculated for each population (Waples & Do 2010).

We defined population census size (Nc) as the total number of adults (over-yearlings) present in the population at the time of sampling (Fall for WB; Summer for FG). This definition was selected because i) YOY abundances are highly variable from year-to-year compared to adult abundances, ii) the proportion of YOY contributing to reproduction is typically close to zero, and iii) this was the definition recommended by Frankham (1995) in his seminal paper and enabled direct comparison with his results. For both populations, the final Nc value was calculated as the number of adults captured in a sample divided by the estimated capture probability of the population (WB = 0.6, FG = 0.65).

To calculate N_e/N_c ratios, we followed guidelines from Waples (2005). Given that brook trout populations display variable age-at-maturity, we adopted comparison strategies proposed for the 'salmon model' (Table 6 within Waples (2005)). Single-sample N_e estimates were divided by N_c values for the same sampling occasion. Temporal N_e estimates were divided by the harmonic mean of N_c values over the

sampling interval. For both methods, harmonic mean N_e estimates of all sampling occasions were divided by the harmonic mean of N_c values over the entire study period.

Results

Genetic Diversity and Genotyping Error

For the WB populations, a total of 1,871 individuals belonging to the 2005 cohort or earlier were genotyped. Loci summary statistics for cohorts belonging to the OS and OL populations are shown in Table 3.1. The only locus testing significant for a heterozygote deficiency was SSaD237 in the 2002 cohort of the OL population. Since this locus was not significant for the remaining three cohorts in this population, and given that observed (0.76) and expected (0.80) heterozygosities did not substantially differ, a null allele was assumed not to be present.

For the FG population, a total of 2,379 individuals captured during the 2004, 2005, and 2006 samples were genotyped. Loci summary statistics for cohorts are shown in Table 3.1. Heterozygote deficiency tests resulted in significant departures from Hardy-Weinberg equilibrium for the SsaD237 locus in both cohorts. The Null allele frequency was estimated to be 0.288 (2004), and 0.300 (2006). The weighted average (0.291) was used to parameterize simulations for pedigree reconstruction accuracy assessment.

Estimated genotyping error rates were low for both populations. For the WB population, complete genotypes were obtained for 91 of the 100 randomly selected individuals. Of these 91 individuals, four contained allele discrepancies between the pair

of genotypes resulting in seven differing alleles. A single individual accounted for four of the differing alleles suggesting a process error for that individual. The resulting per allele error rate was 0.32% (7/2184). For the FG population, a total of 74 PIT tag implanted individuals were captured on two sampling occasions allowing for direct comparison between the two genotypes. Of these, only one individual had a discrepancy between the two generated genotypes, a single miscalled allele. This resulted in a per allele error rate of 0.08% (1/1184).

Pedigree Reconstruction

Sibship reconstruction and parentage assignment analyses performed on the synthetic datasets both indicated a high degree of power to reconstruct full-sibling families and assign parents accurately for both genetic panels. For reconstructed full-sibling families composed of at least two individuals, inferred families had a correct partition rate of 91.2% (0.7%) (SE) (WB) and 95.2% (0.5%) (FG), and assigned parents had an accuracy of 94.2% (0.6%) (WB) and 92.8% (0.9%) (FG). Accuracies for both methods improved as full-sibling family size increased. For example, full-sibling families composed of at least five individuals resulted in accuracies of 97.7% (0.4%) (WB) and 97.1% (0.5) (FG) (sibship), and 96.1% (0.5%) (WB) and 94.8% (1.1%) (FG) (parentage).

As an additional validation of parentage assignment accuracy in the WB populations, known locations of parents during spawning were compared to natal rivers of assigned families for congruence. Of 101 assigned parents available for capture, 84 were detected during the spawning period that produced their assigned family. Of these

84, 76 were captured in the natal river of the assigned family, resulting in a congruence rate of 90.5%.

The number of full-sibling families comprising each cohort varied for both populations (range 28-109 (WB); 24-151 (FG)), however the average family size was relatively stable (4.2 (range 3.9-4.7) (WB); 5.4 (range 3.1-7.6) (FG)) given there was a four-fold (WB) and fifteen-fold (FG) difference in offspring number between the lowest and highest cohort sizes (Table 3.2). The average percentage of assigned parents for full-sibling families was 25% (range 10-38%), while the average percentage of offspring with an assigned parent was 48% (range 17-65%). The average size of full-sibling families with at least one assigned parent were 5.9 (range 4.5-7.2) (WB) and 5.2 (range 2.6-7.9) (FG), indicating that families with assigned parents were representative of the entire family size distribution.

For all cohorts in both populations, the distribution of full-sibling family sizes was highly skewed (Figure 3.2). On average, 16% (range 8-20%) (WB) and 21% (range 16-25%) (FG) of the largest families accounted for 50% of the total number of offspring in a cohort. Alternatively, families composed of only one or two individuals comprised 65% (range 63-68%) (WB) and 64% (range 53-75%) (FG) of the total number of families in a cohort. Thus, reproductive variance among families was high.

The majority of parents in both populations contributed to only one full-sibling family (83% (WB), 77% (FG) (Figure 3.3). For the WB population, where sexes were known for many individuals, females never contributed to more than two full-sibling families, while males accounted for all individuals contributing to three or more families (4%). For both populations, there was a trend for larger parents to contribute to multiple

full-sibling families (Figure 3.4). This in turn translated into larger parents generally contributing more offspring (Figure 3.5). This makes sense for females, since fecundity is a function of size (Letcher *et al.* 2007;Vladykov 1956), however for males it implies that they are mating with larger females.

Examination of the body size of mating-pairs supported the fact that large males were mating with large females, and in general that mating-pairs were size-assortative (Figure 3.6). Isolating for primary mating-pairs (largest full-sibling family produced by a female (WB) or an individual (FG)) substantially increased the fit in the FG population, indicating that secondary mating-pairs were of unequal body size. By itself, body size accounted for 35% (WB) and 39% (FG) of the variation in primary mating- pairs.

Census and Effective Population Size Estimation

The harmonic mean of the adult census size for the FG population was almost thirteen fold greater than that of the OS population, and over three fold greater than that of the OL population (Table 3.3). Not surprisingly, the N_e estimates followed a similar pattern with the FG population having the highest value, followed by the OL and then OS populations (Table 3.3). These rankings were consistent when calculated using either the single-sample or the temporal-sample method.

Estimates produced using the pseudo-maximum likelihood (ML) method, were on average 28% (range 17-34%) higher than those produced using the linkage-disequilibrium (LD) method (Table 3.3). The lower LD estimates were potentially caused by the sample containing a greater proportion of YOY compared to adults given that

brook trout have Type III survivorship (Letcher *et al.* 2007). Thus the estimate would fall between the effective number of breeders (N_b) that produced the YOY cohort and the N_e of the generation (Luikart *et al.* 2010). Alternatively, the pseudo-ML method has been reported to bias N_e estimates high when alleles are present in low frequencies (Jorde & Ryman 2007), which is the case with most microsatellite loci.

N_e estimates produced by the moment method were in better agreement with LD and ML estimates for the OS population than for the OL (Table 3.3). This is not too surprising given that the drift signal is much stronger when the N_e is less than 50 (Luikart *et al.* 2010). The moment N_e estimate for the OL population was most likely biased high by the small sample size of the 2005 cohort (S = 51), resulting in a large proportion of the drift signal being accounted for by the sample size correction factor (1/(2*S)) (Jorde & Ryman 2007). The same explanation holds for the infinite estimate generated by the LD method for the OS 2002 sample (S = 9), where the entire drift signal was accounted for by the correction factor (Waples & Do 2010). Additionally, any error in the life-table data used to calculate the generational overlap correction factor (C) (Jorde & Ryman 1995), would in turn bias the N_e estimate.

Given the potential for bias in all three methods, the logical solution was to average across the estimates by taking their harmonic mean ($\bar{N}_e$) (strategy 1 in Waples & Do (2010)). In spite of the fact that single-sample and temporal-sample methods estimate different N_e (inbreeding (N_eI) versus variance (N_eV), Box 2 in Luikart *et al* (2010)), averaging was justified because single-sample estimates were available for all but one generation (FG 2005) over the time periods used for the temporal estimates (Waples & Do 2010). The resulting $\bar{N}_e$ estimates were 20.6 (WB OS), 113.5 (WB OL), and 187.0

(FG) (Table 3.3). The ratio of effective to census population sizes using these values resulted in the OL population having the highest ratio (0.68), and the FG population having the lowest (0.33) (Table 3.3).

Discussion

The pattern of primary mate choice within these brook trout populations was size-assortative, with 80% of all successful parents contributing only one family, and N_e/N_c greater than four times the average reported in the literature (Frankham 1995;Palstra & Ruzzante 2008). Crespi (1989) hypothesized that size-assortative mating should occur when "large males, large females, or large individuals of both sexes choose large mates because they benefit reproductively and are differentially capable of exercising choice". For salmonids, it makes sense for males to choose larger females because fecundity is positively associated with size (Letcher *et al.* 2007;Vladykov 1956). For females, selection of larger males may have more to do with reduced levels of egg cannibalism than genetic fitness benefits. Blanchfield and Ridgway (1999) reported that female brook trout spawning with relatively smaller males had significantly more eggs eaten by peripheral fish than those spawning with larger males. Furthermore, females in that population were more likely to delay spawning when paired with a relatively smaller male, a behavior also reported for Pacific salmon (Foote 1989;Foote & Larkin 1988). Evidence discounting the genetic quality of large males, or "good genes" hypothesis, was reported by Jacob *et al*. (2007), who concluded that groups of brown trout fertilized by

dominant (large) and non-dominant (small) males did not differ in embryo or juvenile survival.

Given these data, it's apparent that benefits are gained by both sexes through selection of a larger mate. Because the size distribution of males competing for a female should have an upper limit close to the female's size, and because females should choose the largest male, mating pairs would be expected to be of similar size, and thus assortative over the range of fish lengths. What is not apparent is why large males do not attempt to monopolize a range of relatively smaller females, thus increasing their fitness through a higher actualized potential reproduction rate (PRR) (Clutton-Brock 2007). One possible explanation comes not from the perspective of the individual, but from the perspective of the population. Populations following a size-assortative mating pattern would experience lower reproductive variance, and thus maintain greater levels of genetic variation (Wade & Arnold 1980). This would decrease a population's susceptibility to stochastic processes (Newman & Pilson 1997;Saccheri *et al.*1998), and changes in environmental conditions (Frankham *et al.*1999). Indeed, polyandry, a mating strategy leading to reduced reproductive variance, has been hypothesized to have evolved as a mechanism to reduce population extinction risk (Haig & Bergstrom 1995). Furthermore, Price *et al*. (2010) demonstrated its ability to do so in laboratory populations of fruit flies.

Regardless of the underlying cause of size-assortative mating, the end result is an increased proportion of adults contributing to reproduction compared to a dominant pattern (Wade & Arnold 1980). For these brook trout populations, this was evidenced by the high percentage of parents (80%) contributing only one family. Of the remaining 20%, the trend was for larger individuals to contribute to multiple families for both sexes.

For large females, splitting increased fecundity among multiple mates would result in decreased reproductive variance. For males, mating with multiple females would result in increased reproductive variance. Given that contribution to multiple families was approximately equal between the two sexes, any differing effects on reproductive variance should cancel each other out. Instead, the major determinant of reproductive variance within these populations was most-likely the constraint of a female's body size on fecundity, directly limiting offspring number for both her and her mate. This was evidenced by the positive correlation between total number of offspring and parent length, and the skew in family size towards smaller values. However, given that offspring from a large number of families survived the first few months post-emergence, the ontogenetic period that experiences the greatest rate of mortality in salmonids (Einum & Fleming 2000b;Elliott 1984;Letcher *et al.* 2007), and the fact that family size ranks are not maintained through time (Hudy *et al.* 2010), suggests that the fecundity constraint should not impact N_e substantially.

The mean N_e/N_c value for these brook trout populations was four to five times greater than mean values reported for 165 (0.14) (Palstra & Ruzzante 2008) and 102 (0.10) (Frankham 1995) different species. Thus, compared to most species, brook trout in headwater systems have a greater proportion of the population contributing to reproduction through time. High N_e/N_c values have also been reported for other salmonid populations (Araki *et al.* 2007;Ardren & Kapuscinski 2003;Fraser *et al*. 2007), with a trend for higher values to occur in populations of smaller census size (Palstra & Ruzzante 2008). This trend, termed genetic compensation, has been theoretically demonstrated to occur when reproductive variance positively correlates with population census size

(Hedrick 2005). In salmonids with anadromous life-histories, genetic compensation comes from increased reproductive success of mature male parr when densities of adult males are reduced (Jones & Hutchings 2001;Jones & Hutchings 2002). However, for resident populations without this life-history option, a mechanism like size-assortative mating could serve a similar function. More research is needed to assess if this mechanism is present in other headwater stream species, which would suggest an adaptive response by populations inhabiting these highly stochastic environments.

The effect of size-assortative mating on the conservation of genetic variation was evidenced in these brook trout populations by the increase in N_e estimates over the measured time period. This is particularly relevant for the FG population, which was repopulated using only 91 individuals back in 1993 (Hudy *et al.* 2000), and is in complete geographic isolation from other brook trout populations (nearest neighbor is 85 stream km away). The trend of increasing N_e estimates over the sampling period in the two WB populations could potentially be attributed to immigrants from the main-stem population. However, N_e estimates over the same time period for an isolated tributary located in the WB drainage (Letcher *et al.* 2007) resulted in the same pattern (J. Coombs, unpublished data). This suggests that the relatively low N_e estimates for 2002 are better explained by harsh environmental conditions experienced that year, which resulted in reduced survival of potential spawners. The rebound in N_e over the next three years, when environmental conditions were improved and all age-classes contributed to reproduction (J. Coombs, unpublished data), could be explained by genetic compensation from smaller males pairing with smaller females.

The maintenance of N_e in these small, isolated populations suggests that extirpation as a result of inbreeding depression may not be of immediate concern. This conclusion is drawn from the persistence of the FG population despite a founder effect, the persistence of the OS population despite an N_e less than 50 (Franklin & Frankham 1998), and the persistence of the isolated WB population despite an estimated isolation time of 900 years (Letcher *et al.* 2007) and a reduction in heterozygosity of 31% compared to WB populations (0.43 vs 0.63, $N_e = 90$, J. Coombs, unpublished data). Instead, given that these populations have lost genetic variation, and assuming that they are locally adapted to their current environment, we argue that the greater threat to persistence for these populations will arise from an inability to adapt to predicted future changes in environmental conditions (Marshall & Randhir 2008). However, this conclusion depends on habitat area remaining constant in the interim.

Based on these data, strategies for conservation and management would be best served by maintaining or improving connectivity among populations to impede loss of genetic variation, and thus adaptive potential. In a review by Palstra and Ruzzante (2008), open populations had higher heterozygosities and lower genetic diversity loss rates compared to isolated populations of the same N_e. If management actions include population supplementation through hatchery fish, our data suggest the importance of having individuals present over a wide range of lengths to allow for size-assortative mating. This mechanism has the potential to increase recruitment, and enable genetic compensation. The ideal strategy would be to stock individuals at an early life-stage, allowing for increased stocking densities, thus potential for increased genetic diversity, and for survival to occur through natural selection. Additionally, individuals within the

life-stage should have a length distribution reflecting that found in the wild population, as individual length ranks have been shown to be highly stable throughout a cohort's lifespan (Letcher *et al*. 2010).

In summary, brook trout populations inhabiting headwater streams follow a size-assortative mating pattern. This resulted in reduced reproductive variance among individuals, as evidenced by the large number of individuals contributing to only a single family, which in turn produced a relatively high N_e/N_c ratio. We hypothesize that size-assortative mating within these populations has evolved as a mechanism to respond to population fluctuations caused by stochastic processes. Such a mechanism acts to conserve genetic variation, and thus reduce local extinction probability. From a conservation and management perspective, we argue that brook trout populations are more susceptible to an inability to adapt to environmental change than from inbreeding depression. We recommend maintaining or improving population connectivity to buffer against loss of genetic diversity, and, if supplemental strategies must be used, ones that result in length distributions observed in wild populations. Our results highlight the importance of understanding mating decisions in order to improve conservation and management strategies in the face of continued anthropogenic impacts and looming climate change.

Table 3.1. Single locus summary statistics for West Brook (WB) and Fridley Gap (FG) brook trout cohorts. Measures for each locus are: (A_O) observed number of alleles; (H_O) observed heterozygosity; (H_E) expected heterozygosity; (f_{is}) an analogue of Wright's F_{IS} statistic; (p) probability of departure from Hardy-Weinberg expectations in the direction of heterozygote deficiency. Bold p values indicate significant genotypic departures from expected Hardy-Weinberg equilibrium when evaluated using a sequential Bonferroni correction for multiple tests (WB: k = 48, α = 0.05; FG: k = 16, α = 0.05).

Drainage	Population	Cohort		SfoC24	SfoC88	SfoC113	SfoC115	SfoC129	SfoD75	SfoD100
WB	Open-Small	2002	A_O	4	3	7	4	4	7	5
		(n=54)	H_O	0.26	0.43	0.96	0.46	0.41	0.69	0.46
			H_E	0.28	0.36	0.79	0.52	0.40	0.71	0.57
			f_{is}	0.08	-0.18	-0.22	0.12	-0.02	0.04	0.19
			p	0.49	1.00	1.00	0.03	0.49	0.40	0.15
		2003	A_O	3	4	5	4	5	7	4
		(n=40)	H_O	0.78	0.45	0.53	0.9	0.88	0.85	0.85
			H_E	0.66	0.53	0.49	0.52	0.81	0.71	0.62
			f_{is}	-0.17	0.14	-0.06	-0.74	-0.08	-0.20	-0.39
			p	0.96	0.17	0.40	1.00	0.80	0.99	1.00
		2004	A_O	4	4	8	6	5	7	5
		(n=201)	H_O	0.35	0.39	0.82	0.77	0.72	0.87	0.45
			H_E	0.31	0.39	0.79	0.70	0.66	0.75	0.58
			f_{is}	-0.14	0.00	-0.04	-0.09	-0.09	-0.16	0.22
			p	1.00	0.98	0.87	1.00	0.97	1.00	0.08
		2005	A_O	4	3	7	6	5	8	5
		(n=73)	H_O	0.34	0.66	0.78	0.64	0.51	0.73	0.77
			H_E	0.32	0.50	0.78	0.56	0.51	0.67	0.65
			f_{is}	-0.06	-0.32	0.00	-0.15	0.00	-0.08	-0.18
			p	0.70	1.00	0.64	0.99	0.67	0.88	1.00

Table 3.1. (continued)

Drainage	Population	Cohort		SsaD237	SfoB52	SfoC38	SfoC86	SfoD91a	Average
WB	Open-Small	2002	A_O	9	3	3	2	4	4.58
		(n=54)	H_O	0.87	0.65	0.20	0.41	0.81	0.55
			H_E	0.74	0.58	0.22	0.47	0.59	0.52
			f_{is}	-0.18	-0.12	0.06	0.14	-0.39	-0.04
			p	1.00	0.97	0.45	0.24	1.00	
		2003	A_O	9	3	3	2	5	4.5
		(n=40)	H_O	0.83	0.63	0.45	0.45	0.65	0.69
			H_E	0.75	0.66	0.39	0.38	0.55	0.59
			f_{is}	-0.10	0.05	-0.17	-0.19	-0.19	-0.17
			p	0.60	0.36	0.93	0.96	0.96	
		2004	A_O	13	4	3	4	7	5.83
		(n=201)	H_O	0.91	0.74	0.40	0.46	0.71	0.63
			H_E	0.84	0.61	0.46	0.40	0.65	0.59
			f_{is}	-0.08	-0.21	0.13	-0.15	-0.10	-0.06
			p	0.99	1.00	0.17	0.99	0.29	
		2005	A_O	10	5	3	4	7	5.58
		(n=73)	H_O	0.92	0.52	0.25	0.53	0.67	0.61
			H_E	0.84	0.50	0.22	0.45	0.67	0.56
			f_{is}	-0.09	-0.05	-0.10	-0.19	0.00	-0.10
			p	0.43	0.79	1.00	0.98	0.66	

Table 3.1. (continued)

Drainage	Population	Cohort		SfoC24	SfoC88	SfoC113	SfoC115	SfoC129	SfoD75	SfoD100
WB	Open-Large	2002	A_O	4	5	8	9	5	9	6
		(n=161)	H_O	0.13	0.6	0.71	0.36	0.55	0.74	0.62
			H_E	0.12	0.58	0.74	0.37	0.50	0.78	0.64
			f_{is}	-0.05	-0.04	0.03	0.03	-0.11	0.05	0.03
			p	1.00	0.86	0.27	0.01	0.94	0.01	0.58
		2003	A_O	4	5	8	10	5	11	6
		(n=208)	H_O	0.15	0.56	0.81	0.34	0.46	0.75	0.63
			H_E	0.16	0.54	0.77	0.38	0.47	0.78	0.65
			f_{is}	0.06	-0.03	-0.05	0.13	0.03	0.04	0.04
			p	0.24	0.88	0.96	0.26	0.34	0.03	0.57
		2004	A_O	4	4	9	9	5	10	8
		(n=185)	H_O	0.08	0.51	0.71	0.22	0.48	0.77	0.62
			H_E	0.08	0.52	0.75	0.25	0.46	0.80	0.61
			f_{is}	-0.03	0.02	0.05	0.13	-0.04	0.05	-0.01
			p	1.00	0.56	0.19	0.53	0.17	0.18	0.15
		2005	A_O	3	4	7	7	5	7	5
		(n=51)	H_O	0.16	0.55	0.67	0.22	0.45	0.78	0.61
			H_E	0.15	0.56	0.76	0.25	0.48	0.80	0.58
			f_{is}	-0.06	0.02	0.12	0.15	0.05	0.02	-0.04
			p	1.00	0.47	0.13	0.08	0.10	0.49	0.60

Table 3.1. (continued)

Drainage	Population	Cohort		SsaD237	SfoB52	SfoC38	SfoC86	SfoD91a	Average
WB	Open-Large	2002	A_O	18	6	3	6	9	7.33
		(n=161)	H_O	0.76	0.68	0.61	0.39	0.73	0.57
			H_E	0.80	0.65	0.51	0.41	0.77	0.57
			f_{is}	0.05	-0.04	-0.19	0.05	0.05	-0.01
			p	**0.00**	0.84	0.96	0.46	0.17	
		2003	A_O	24	7	3	5	10	8.17
		(n=208)	H_O	0.88	0.66	0.61	0.51	0.65	0.58
			H_E	0.89	0.71	0.53	0.50	0.75	0.60
			f_{is}	0.01	0.06	-0.14	-0.03	0.13	0.02
			p	0.63	0.30	1.00	0.79	0.01	
		2004	A_O	22	7	3	5	10	8
		(n=185)	H_O	0.81	0.72	0.36	0.52	0.72	0.54
			H_E	0.85	0.74	0.40	0.50	0.74	0.56
			f_{is}	0.04	0.02	0.10	-0.05	0.02	0.02
			p	0.08	0.52	0.08	0.77	0.05	
		2005	A_O	16	5	3	5	6	6.08
		(n=51)	H_O	0.71	0.71	0.49	0.57	0.76	0.56
			H_E	0.74	0.71	0.45	0.52	0.72	0.56
			f_{is}	0.05	0.01	-0.09	-0.10	-0.06	0.01
			p	0.44	0.57	0.84	0.85	0.82	

Table 3.1. (continued)

Drainage	Population	Cohort		SfoC24	SfoC88	SfoC113	SfoC115	SfoC129	SfoD75	SfoD100
FG	-	2004	A_O	6	7	11	18	4	11	12
		(n=899)	H_O	0.71	0.79	0.83	0.81	0.66	0.87	0.83
			H_E	0.70	0.75	0.81	0.85	0.64	0.84	0.85
			f_{is}	-0.05	-0.02	-0.03	0.05	0.03	0.03	-0.02
			p	0.92	0.71	0.99	0.14	0.05	0.32	0.64
		2006	A_O	6	6	10	15	4	9	11
		(n=104)	H_O	0.68	0.66	0.83	0.89	0.65	0.84	0.86
			H_E	0.74	0.72	0.73	0.85	0.65	0.86	0.84
			f_{is}	0.03	0.04	0.01	0.00	-0.04	0.04	-0.03
			p	0.09	0.30	0.16	0.42	0.66	0.05	0.90

Table 3.1. (continued)

Drainage	Population	Cohort		SsaD237	SfoB52	SfoC38	SfoC86	SfoD91a	Average
FG	-	2004	A_O	18					10.88
		(n=899)	H_O	0.46					0.74
			H_E	0.87					0.79
			f_{is}	0.51					0.06
			p	**0.00**					
		2006	A_O	16					9.63
		(n=104)	H_O	0.36					0.72
			H_E	0.84					0.78
			f_{is}	0.52					0.07
			p	**0.00**					

Table 3.2. Descriptive measures for reconstructed full-sibling families containing a minimum of two individuals for brook trout populations inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. N_O = Number of offspring, N_F = Number of full-sibling families, N_P = Number of assigned parents, N_{PP} Number of full-sibling families with both parents assigned, PO = Proportion of offspring assigned a parent.

Drainage	**Cohort**	$\mathbf{N_O}$	$\mathbf{N_F}$	$\mathbf{N_P}$	$\mathbf{N_{PP}}$	**PO**
WB	2002	426	109	30	2	0.30
	2003	298	71	38	11	0.61
	2004	401	84	44	11	0.60
	2005	112	28	21	7	0.65
FG	2004	1154	151	108	25	0.57
	2006	75	24	5	0	0.17

Table 3.3. Estimates of effective population size (N_e) generated by single-sample and temporal methods for brook trout populations inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. N_c = Estimated census size of adult (over-yearling) brook trout; S = Sample size of individual genotypes used to generate N_e estimate; N_e = Estimate of effective population size; $\bar{N}_e$ = Harmonic mean of N_e over all methods; INF = Infinite N_e estimate; N/A = Calculation not applicable; - = Calculation not available.

				Single-Sample[1]			Temporal Sample						
								ML[2]		Moment[3]			
Drainage	**Population**	**Cohort**	**N_c**	**S**	**N_e**	**N_e/N_c**	**S**	**N_e**	**N_e/N_c**	**N_e**	**N_e/N_c**	**$\bar{N}_e$**	**$\bar{N}_e/N_c$**
WB	OS	2002	20	9	INF	N/A	54	-	-	-	-		
		2003	68	66	11.7 (9,14)	0.17	40	13.4 (11,16)	0.43	12.0 (8,17)	0.39		
		2004	58	100	19.7 (17,23)	0.34	201	26.8 (22,33)	0.43	17.5 (12,24)	0.28		
		2005	115	64	32.1 (24,42)	0.28	73	30.9 (26,36)	0.40	28.5 (20,39)	0.37		
		Mean[4]	44		19	0.43		28.9	0.65	18.5	0.39	20.6	0.47

[1] Calculated using linkage disequilibrium (Waples & Do, 2008).

[2] Calculated using pseudo-maximum likelihood (Wang, 2001).

[3] Calculated using unbiased F estimator (Jorde & Ryman, 2007) and generational overlap correction factor (Jorde & Ryman, 1995).

[4] N_c mean is the harmonic mean; Single-sample mean is the weighted harmonic mean; ML temporal mean is the N_e estimate using all cohorts simultaneously; Moment temporal mean is the harmonic mean of each interval estimate.

[5] N_c mean includes adults captured during the 2005 sample

Table 3.3. (continued)

				Single-Sample[1]			Temporal Sample						
								ML[2]		Moment[3]			
Drainage	**Population**	**Cohort**	**N_c**	**S**	**N_e**	**N_e/N_c**	**S**	**N_e**	**N_e/N_c**	**N_e**	**N_e/N_c**	**$\bar{N}_e$**	**$\bar{N}_e/N_c$**
WB	OL	2002	137	86	59.3 (44,81)	0.43	161	-	-	-	-		
		2003	183	211	87.4 (67,115)	0.48	208	100.4 (79,134)	0.64	105.7 (78,138)	0.67		
		2004	143	122	126.8 (101,164)	0.88	185	101.0 (80, 133)	0.63	118.7 (87,155)	0.74		
		2005	228	115	140.3 (105,197)	0.61	51	178.1 (96,1012)	1.01	224.2 (163,295)	1.27		
		Mean[4]	166		96.6	0.58		116	0.70	134.2	0.81	113.5	0.68

[1] Calculated using linkage disequilibrium (Waples & Do, 2008).

[2] Calculated using pseudo-maximum likelihood (Wang, 2001).

[3] Calculated using unbiased F estimator (Jorde & Ryman, 2007) and generational overlap correction factor (Jorde & Ryman, 1995).

[4] N_c mean is the harmonic mean; Single-sample mean is the weighted harmonic mean; ML temporal mean is the N_e estimate using all cohorts simultaneously; Moment temporal mean is the harmonic mean of each interval estimate.

[5] N_c mean includes adults captured during the 2005 sample

Table 3.3. (continued)

				Single-Sample[1]			Temporal Sample						
								ML[2]		Moment[3]			
Drainage	**Population**	**Cohort**	**N_c**	**S**	**N_e**	**N_e/N_c**	**S**	**N_e**	**N_e/N_c**	**N_e**	**N_e/N_c**	**$\bar{N}_e$**	**$\bar{N}_e/N_c$**
FG	-	2004	635	1254.2	148.2 (128,171)	0.23	1190	-	-	-	-		
		2006	750	533.7	168.8 (147,195)	0.23	101.4	235.5 (175,359)	0.42	-	-		
		Mean[4]	5675		155	0.27		235.5	0.42	-	-	187	0.33

[1] Calculated using linkage disequilibrium (Waples & Do, 2008).

[2] Calculated using pseudo-maximum likelihood (Wang, 2001).

[3] Calculated using unbiased F estimator (Jorde & Ryman, 2007) and generational overlap correction factor (Jorde & Ryman, 1995).

[4] N_c mean is the harmonic mean; Single-sample mean is the weighted harmonic mean; ML temporal mean is the N_e estimate using all cohorts simultaneously; Moment temporal mean is the harmonic mean of each interval estimate.

[5] N_c mean includes adults captured during the 2005 sample

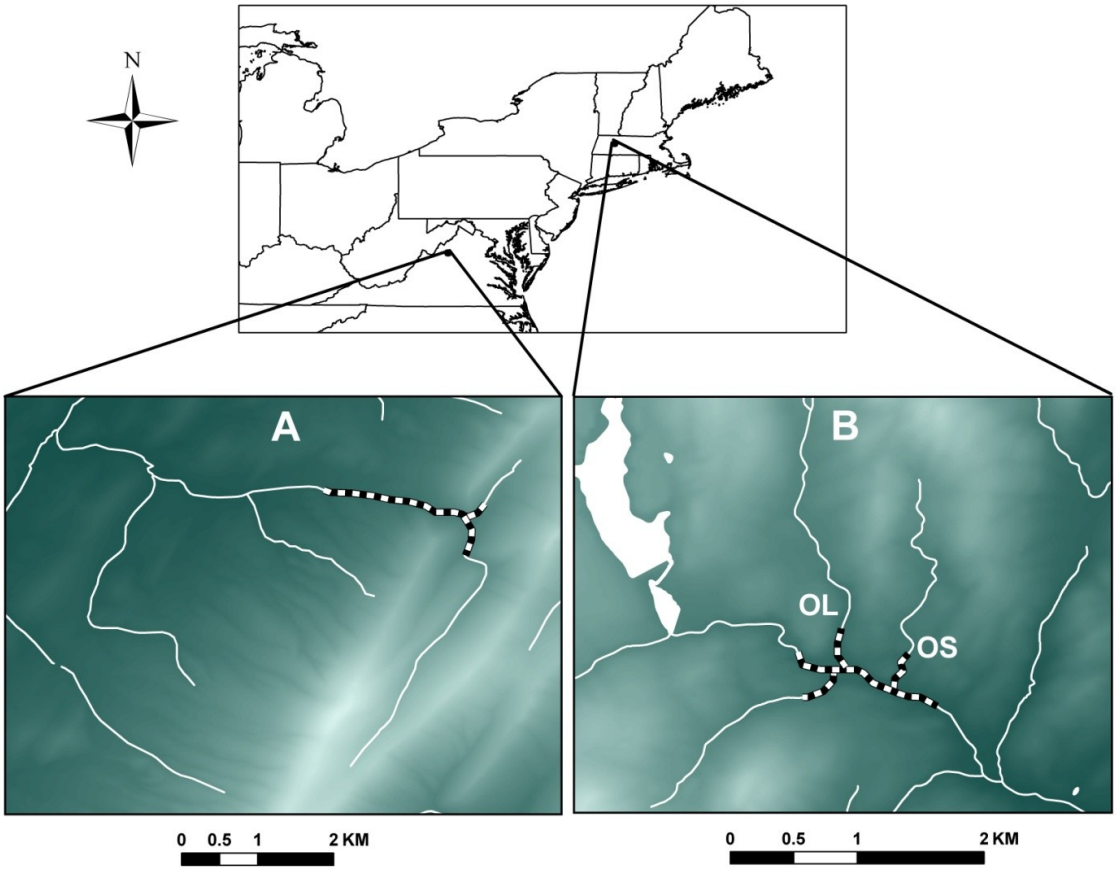


Figure 3.1. Map of study regions. A) Fridey Gap (FG) drainage located in Rockingham County, VA, USA. B) West Brook (WB) drainage located in Whately, MA, USA. OL = Open-Large population; OS = Open-Small population. Dashed line indicates sample reaches.

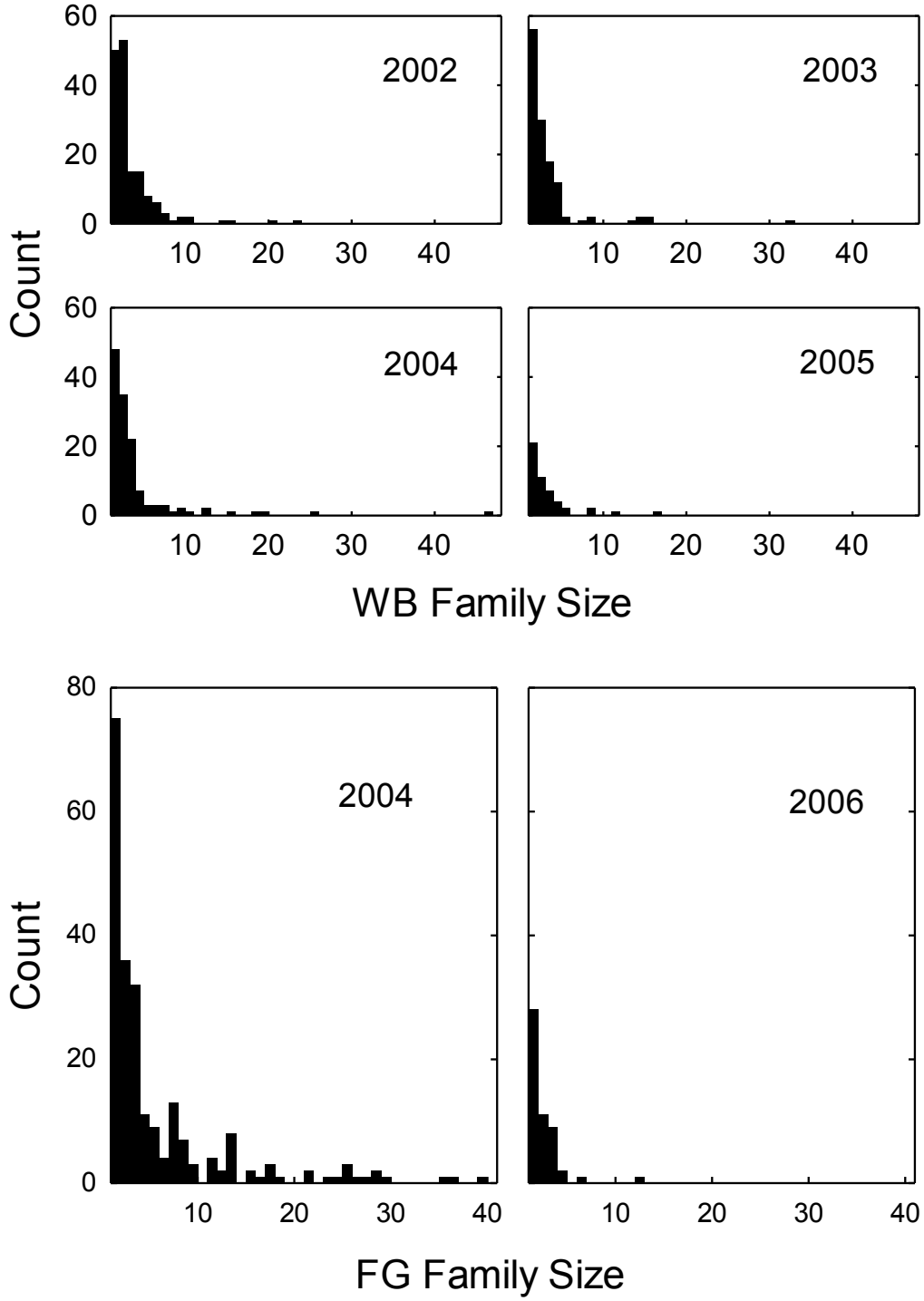


Figure 3.2. Histograms of full-sibling family sizes by cohort for brook trout populations inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. Number of families and number of offspring are as follows: WB: 2002 (159, 476); 2003 (127, 354); 2004 (132, 449); 2005 (49, 133); FG: 2004 (226-1229); 2006 (52, 103).

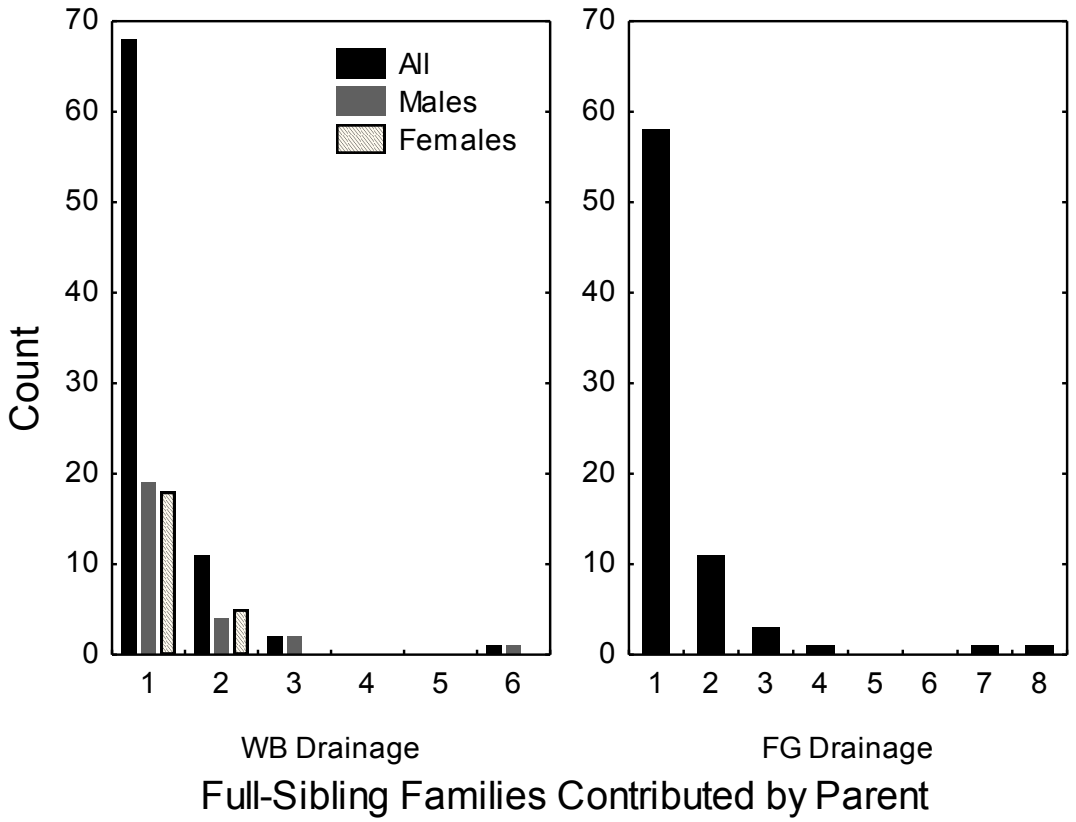


Figure 3.3. Individual contribution of full-sibling families by parents for brook trout inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. The category 'All' includes both sexed and unsexed individuals. Number of parents for WB = 82 (All), 26 (Males), 23 (Females). Number of parents for FG = 75 (All).

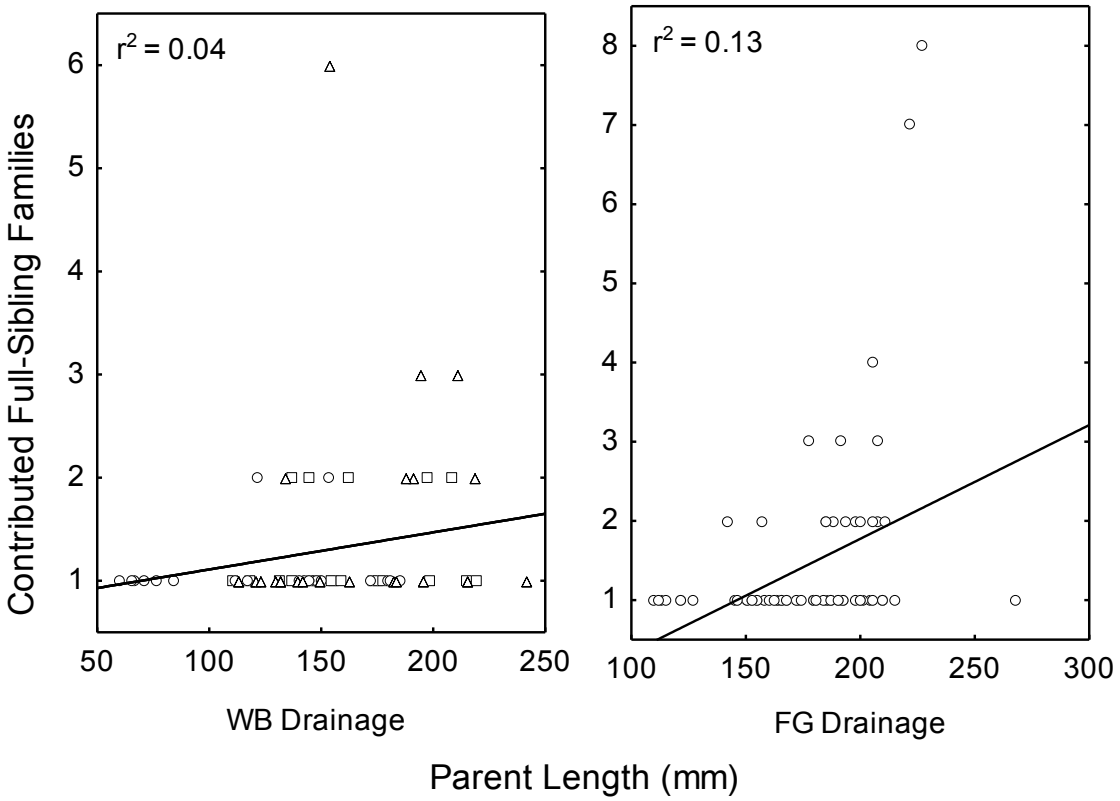


Figure 3.4. Individual full-sibling family contributions as a function of individual length for brook trout populations inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. Squares represent females, triangles represent males, and circles represent individuals of unknown sex. Sample sizes are 67 (WB [Females = 20; Males = 26; Unknown = 21]) and 74 (FG). Lines represent best fit linear regressions for all individuals (WB: Y = 0.7495 + 0.0036*X, p = 0.115; FG: Y = -1.0992 + 0.0144*X, p = 0.002).

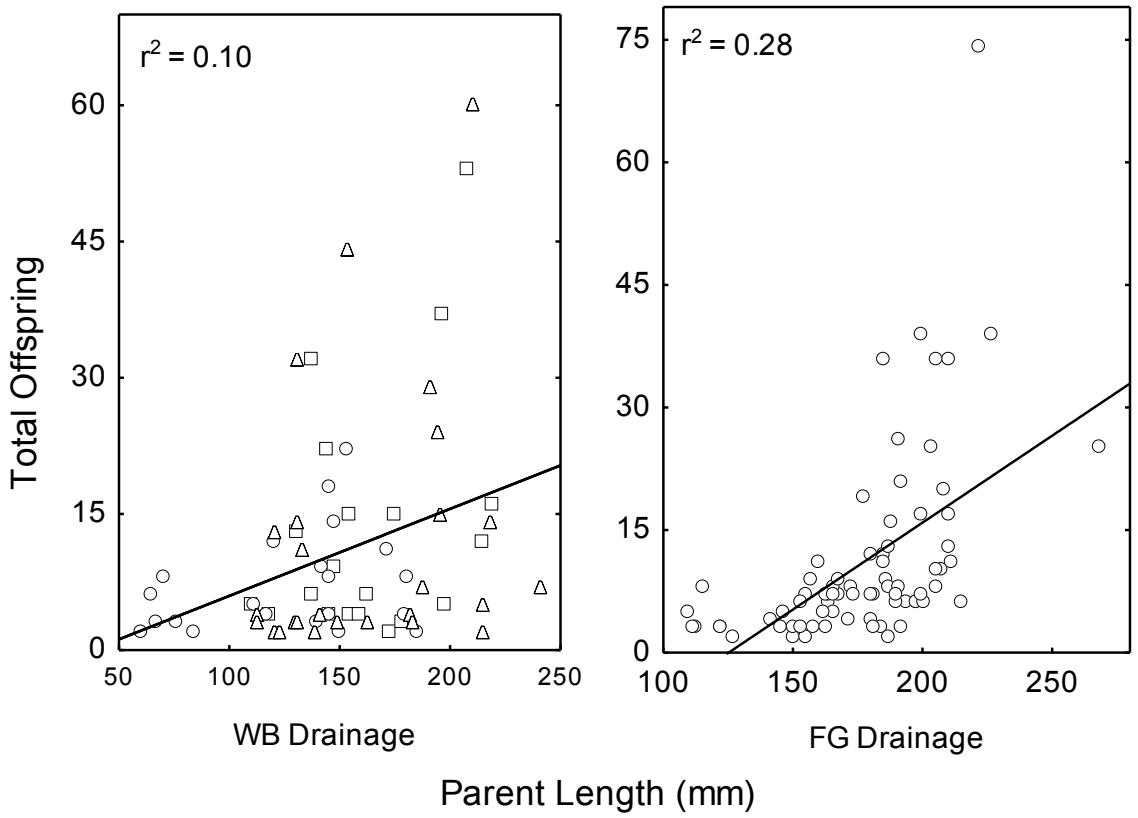


Figure 3.5. Total offspring contributed by an individual as a function of individual length for brook trout populations inhabiting the West Brook (WB) and Fridley Gap (FG) drainages. Squares represent females, triangles represent males, and circles represent individuals of unknown sex. Sample sizes are 67 (WB [Females = 20; Males = 26; Unknown = 21]) and 74 (FG). Lines represent best fit linear regressions for all individuals (WB: Y = -3.6077 + 0.0958*X, p = 0.008; FG: Y = -26.6304 + 0.2127*X, p = 0.000).

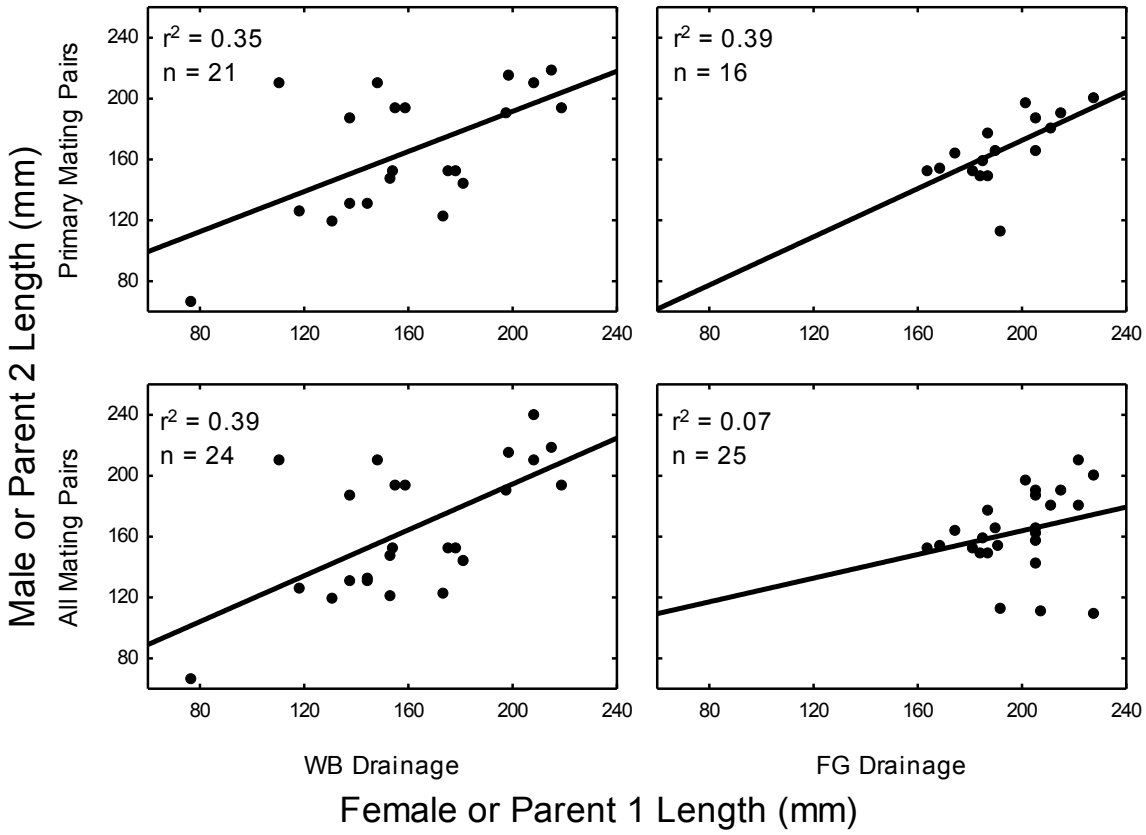


Figure 3.6. Size of mating pairs in the West Brook (WB) and Fridley Gap (FG) brook trout populations. Individuals in the WB population were of known sex, while those in the FG population were unsexed. For sexed individuals, 'primary mating pair' refers to the largest full-sibling family produced by a female. For unsexed individuals, 'primary mating pair' refers to the largest full-sibling family produced by an individual. 'All mating pairs' includes both primary and non-primary mating pairs for individuals reproducing multiple times. Lines represent best fit linear regressions (WB Primary: Y = 43.787 + 0.7538 * X, $p = 0.005$; WB All: Y = 59.8417 + 0.6587 * X, $p = 0.001$; FG Primary: Y = 14.0473 + 0.7923 * X, $p = 0.010$; FG All: Y = 85.998 + 0.3891 * X, $p = 0.214$).

CHAPTER IV

WHEN TO REPRODUCE: CAUSES AND CONSEQUENCES OF EARLY MATURATION IN AN INDETERMINATELY GROWING SPECIES

Abstract

Life history theory predicts that age at maturity is determined by tradeoffs between benefits gained from early reproduction with costs imposed on future reproduction. For a species where female fecundity is positively correlated with size, and mating pairs are size-assortative, an increased size-at-age would yield increased fitness, and thus should shift the balance towards a younger age-at-maturity. We assessed this hypothesis by evaluating the causes and consequences of age-specific maturity for three cohorts of a wild brook trout population. Individual age-at-maturity ranged from 0 to 2 years, with the proportion of age-0 and age-1 individuals maturing in a given year dependent upon growth opportunities determined primarily by environmental conditions. Mature fish were significantly larger than immature fish within an age-class, however, their survival rates were similar. Furthermore, parental length did not influence offspring survival. These data suggest that the cost of early maturation is instead manifested through a reduction in egg number for females, and a reduced ability to acquire mates for males, both determined by an individual's size. Indeed, fecundity predicted by mean length of immature and mature fish within an age-class would result in mature fish producing an average of 38% (age-0) and 33% (age-1) more eggs than immature fish. For

this population, the effect of environmental conditions on growth rate was sufficient to shift the balance towards earlier age-at-maturation. Evolutionarily, this mechanism allows the population to counter stochastic processes that decrease survival through increased production potential. However, utility of this mechanism is strongly constrained by environmental conditions.

Introduction

For a population to persist, recruitment must on average exceed losses. In naturally reproducing populations, fluctuations in environmental conditions are typically reflected by fluctuations in population size (Wilson *et al*. 2009). To overcome the added losses experienced during periods of harsh environmental conditions, a population must find a way to increase production. One mechanism that enables populations to achieve this is a decrease in the average age-at-maturity, thus resulting in a greater proportion of the population contributing to reproduction (Cole 1954;Rose *et al*. 2001).

From life-history theory, age-at-maturity is predicted to occur when the benefits gained from maturation exceed the costs (Stearns 1992). Benefits of early maturation include an increased probability of surviving to maturity, and a higher fitness as a result of offspring being able to reproduce sooner (Cole 1954). Costs of early maturation include reduced fecundity, and potentially diminished offspring survival as a consequence of inferior quality or parental care (Stearns 1992). Thus, the decision to mature is based on the tradeoff between current offspring production (fewer number or

poorer quality), and future offspring production (greater number or better quality, but contingent upon survival until the next reproductive bout) (Rose & Charlesworth 1981).

In populations, age-at-maturity is determined by mortality rate, with decreased adult survival selecting for earlier maturation (Harvey & Zammuto 1985;Rose & Charlesworth 1981;Winemiller & Rose 1992). Under constant environmental conditions, selection would result in a single optimum age maximizing fitness (Bell 1980). However, in natural populations both abiotic and biotic conditions vary considerably, and this is reflected in the substantial individual variation seen in age-at-maturation (Stearns 1992). This implies that the costs and benefits of early maturation may differ among individuals, and thus multiple optima exist within a population (Bell 1980;Gadgil & Bossert 1970).

To account for the presence of multiple optima, it was postulated that developmental thresholds must be achieved in order to transition between life-history stages (Wilbur & Collins 1973). This hypothesis was put forth specifically to account for the negative relationship between growth conditions and age at a transition reported for most species (Berrigan & Charnov 1994;Stearns & Koella 1986). Indeed, life history models accounting for developmental thresholds were able to reproduce this pattern (Day & Rowe 2002). Thus mechanisms affecting growth opportunity have the potential to further influence age-at-maturity, and by default, a population's production potential.

To assess the costs and benefits of early maturation for both the individual and the population, we measured age-at-maturity for an intensively studied wild brook trout (*Salvelinus fontinalis*) population. Like most salmonids, brook trout have flexible life-histories as a result of indeterminate growth and size-dependent reproductive strategies (Hendry & Stearns 2004). Reproduction occurs annually during the fall, with female

fecundity positively correlated with length (Vladykov 1956), and mate choice following a size-assortative pattern (Blanchfield & Ridgway 1999). Juvenile survival is low, with the majority of mortality occurring during the first few months post-emergence (Einum & Fleming 2000b;Elliott 1984;Letcher *et al*. 2007), and growth opportunity is principally determined by environmental conditions and population density (Jenkins *et al*. 1999;Nislow *et al.* 2004;Vincenzi *et al*. 2007;Xu *et al.* 2010).

We studied two headwater tributary populations over three successive years during which time size attributes, maturity status, reproductive success, and survival were recorded or determined for a large proportion of individuals in each population. We predicted that under harsh environmental conditions or high densities, age-at-maturity would increase due to decreased growth opportunity, and thus decreased exceedence of the developmental threshold. We further predicted that individuals maturing early within an age-class would be larger than their non-maturing counterparts, resulting in size-related benefits, however, they would also experience increased costs in the form of reduced survival to both themselves and their offspring. To evaluate these predictions, we 1) assessed rates of early maturation under varying environmental conditions and population densities, 2) compared growth metrics between immature and mature fish within an age-class to assess potential developmental thresholds, 3) determined costs and benefits of early maturation in terms of parent and offspring survival and fecundity tradeoffs, and 4) related our findings to population recruitment and persistence.

Materials and Methods

Study Site and Sampling Design

The study was conducted in the West Brook watershed located in Whately, MA, USA (watershed area 11.8 km2) (Figure 4.1). Stream habitat and land use are described in Letcher *et al*. (2002; 2007). The study area was an approximately 1 km stretch of stream with two tributaries entering into the main-stem (WB) at river kilometers 0.4 (Open-Small (OS), watershed area 1.1 km2) and 0.6 (Open-Large (OL), watershed area 2.4 km2). The downstream end of the WB was bounded by a small waterfall (1-m tall, passable by fish), while the upstream end was unbounded. Each tributary study area was 300 m long and were both bounded by impassable waterfalls at their upstream ends. The confluence of the OL tributary was open, but the confluence of the OS tributary was interrupted by a perched culvert (≈1 m tall, passable by fish (Letcher *et al.* 2007)). Average wetted widths of the three streams were 4.5 m (WB), 2 m (OS), and 3 m (OL).

During the study, the streams were inhabited by naturally reproducing brook trout and brown trout (*Salmo trutta*) populations. Additionally, Atlantic salmon (*Salmo salar*) fry (25 mm) were stocked in the WB annually during spring through 2004. The only other fish species consistently present was blacknose dace (*Rhinichthys atratulus*), although abundances were low. Sampling was conducted from June 2002 to December 2006. Samples were seasonal, typically in March, June, September and December, although ice buildup precluded December sampling in the main-stem four of the five years, and in the tributaries one of the five years. A total of 18 sampling occasions were

completed during the study period. Sampling protocol for the main-stem consisted of two-pass electrofishing (300 V unpulsed DC current) with block nets set at the upstream and downstream ends of each 20 m long section. Tributaries were sampled using single-pass electrofishing without block nets. Upon capture, length (± 1 mm fork length), weight (± 0.1 mg wet weight), location (section), and sex and maturity status (if caught during the fall) were recorded for each fish. Untagged fish were implanted with 12-mm PIT tags (Digital Angel, St. Paul MN, USA) if fork length exceeded 60 mm (Gries & Letcher 2002). Cohort was assigned if the fish was captured during its first year (size distribution did not overlap other age-classes), otherwise cohort was assigned based on length distributions of known-aged fish. Anal fin clips for genetic analysis were taken on all untagged fish. Following work-up, fish were returned to their section of capture.

Genetic Diversity and Genotyping Error

Because brook trout are cryptic breeders with no parental care, we used individual genotypes to reconstruct the pedigree structure among sampled individuals. A panel of twelve microsatellite loci (SfoB52, SfoC24, SfoC38, SfoC86, SfoC88, SfoC113, SfoC115, SfoC129, SfoD75, SfoD91a, SfoD100 (King *et al*. 2003), SsaD237 (King *et al.* 2005)) was selected based on its ability to accurately reconstruct full-sibling families and assign parents for simulated data (see below). Protocols for DNA extraction and amplification followed King *et al*. (2005). Loci were electrophoresed on an ABI Prism 3100-Avant genetic analyzer (Applied Biosystems Inc., Foster City, CA), and alleles were scored using GENESCAN v3.7 software (Applied Biosystems Inc.).

Standard measures of genetic diversity were calculated to assess marker quality. Allele number and observed and expected heterozygosities were calculated using GDA v1.0 (Lewis & Zaykin 2001). Estimation of f_{is}, an analogue of Wright's F_{IS}, and testing for departures from Hardy-Weinberg equilibrium were performed using GENEPOP v4.0.10 (Rousset 2008). Tests were conducted using the heterozygote deficiency option to detect the presence of null alleles. Tests were performed on the 2002-2005 cohorts for each locus in each tributary (k = 48), and significance was assessed using a sequential Bonferroni correction (Holm 1979;Rice 1989) with an α of 0.05.

Because genotyping error has the potential to bias pedigree reconstruction (Wilson & Ferguson 2002), the genotyping error rate was assessed and used to minimize its impact. A second DNA extraction and amplification for all twelve loci was performed on 100 randomly selected individuals. Alleles were compared between the two genotypes for each individual, and a per allele error rate estimate was obtained.

Pedigree Reconstruction

The power of the loci panel to accurately reconstruct full-sibling families and assign parents was assessed through the use of synthetic data. Simulated populations were generated using the program PEDAGOG v1.2 (Coombs *et al.* 2010a). Genetic and demographic parameters defining the simulated population were derived using field data from the study population. The simulated population was subjected to a seasonal sampling scheme using field-derived capture probability estimates. Sibship reconstruction and initial parentage assignment analysis were performed on sampled

individuals from the simulated population using the programs COLONY v1.2 (Wang 2004) (sibship) and PEDAPP v1.1 (Almudevar 2007) (parentage). Final parentage assignments were acquired using the sibship constraint (SC) method within the program PEDAGREE v1.04 (Coombs *et al.* 2010b). The SC method was run using a minimum threshold value of 0.2501 for full-sibling families with two members, and 0.1667 for full-sibling families with three or more members. The results from the two runs were then merged. Accuracy of reconstructed families and assigned parents was calculated using PEDAGREE. A total of ten replicates were simulated. The same methodology was used to construct the pedigree for the West Brook dataset.

Environmental Variables

To assess the extent to which environmental conditions influenced age-at-maturation, stream flow and temperature were recorded for each stream. We focused on the spring and summer seasons because the decision to mature has been reported to be dependent upon energy reserves during the spring (Thorpe 1994), and the majority of growth in this system occurs during these seasons (Xu *et al.* 2010). The dates defining seasons (3/1-6/22 (spring), 6/23-10/21 (summer)) were aligned with sampling intervals in order to relate individual size and growth attributes to environmental conditions.

Water temperature (±0.01 oC) was measured every 2 h for each stream over the entire study period using data loggers (Onset Computer Corp, Pocasset, MA, U.S.A.). Average daily temperature was used to calculate seasonal means. Stream flow (m3/s) for the WB was estimated using a flow extension model (Nielsen 1999) based on data from a

nearby USGS stream gage (Mill River, Northampton, MA, U.S.A.). Correlation between estimates from the flow extension model and values based on a stage-discharge relationship developed at the study site (continuous stage record and several direct measurements of discharge) was high (r = 0.94, N = 506). Flow extension estimates were used instead of the stage-discharge values because of data gaps and problems associated with icy conditions. Tributary flows were estimated in the same manner, with the exception that they were related to West Brook flow extension estimates (OL: r = 0.96, N = 100; OS: r = 0.86, N = 90). Seasonal means were calculated from average daily stream flow values.

Population Densities and Length Metrics

Because growth rate can be affected by density-dependent processes (Grossman *et al*. 2010;Xu *et al.* 2010), densities of YOY and adult brook trout were calculated. Values were based on numbers of individuals captured in the fall sample because YOY had not yet emerged during the spring sample, and capture probability for YOY was low during the summer sample. The area of each tributary was calculated by multiplying the length of the sampled reach by the average wetted width based on mean summer flow (OL: Y=0.4379*ln(X)+4.737, r = 0.90, N = 20; OS: Y=0.2441*ln(X)+2.9989, r = 0.78, N = 22). Density (fish/m2) was calculated by dividing the number of YOY or adult brook trout captured in each tributary during the fall sample by the estimated area.

To ascertain the relationship between body size and maturity status, mean lengths were calculated for immature and mature fish within age-classes. A fish was determined

to be mature for a spawning year if it was caught in the fall sample and observed in a mature state, or if it was genetically assigned as a parent to a family produced by that spawning year. To determine cause of maturation, length and growth rate were examined in the previous spring for immature and mature fish. Growth rate was calculated as the change in length divided by the number of days between spring and fall capture. Additionally, condition factor was calculated for both the previous spring and summer to bracket the potential maturation decision period (Thorpe 1994). Values were calculated using the formula Weight*10,000/Length^3.

Survival Analysis

Survival estimates were generated to determine the direct and indirect costs of maturation. The direct cost was measured as the difference in survival between mature and immature groups. Estimates of survival were generated for each age-class, and for all age-classes pooled together. The indirect cost was measured as the effect of parental size on offspring survival. Parental length during spawning was used as an individual covariate for survival estimation. To account for a possible parental sex effect, separate estimates were generated using maternal and paternal lengths. All survival estimates were obtained using the program MARK v5.1 (White & Burnham 1999). To ensure assumptions of the survival model were not violated, goodness-of-fit tests were performed on each dataset under the single-state model using the program U-CARE v2.3.2 (Choquet *et al*. 2009). Significant differences between survival estimates for maturity classes in both pooled and age-specific groups were assessed through

comparison of 95% confidence intervals for the beta parameters. Estimates without overlap were considered significant.

Results

Genetic Diversity and Genotyping Error

A total of 1,871 individuals belonging to the 2005 cohort or earlier were genotyped. Loci summary statistics for cohorts belonging to the OS and OL populations are shown in Table 4.1. The only locus testing significant for a heterozygote deficiency was SSaD237 in the 2002 cohort of the OL population. Since this locus was not significant for the remaining three cohorts in this population, and given that observed (0.76) and expected (0.80) heterozygosities did not substantially differ, a null allele was assumed to not be present.

For genotyping error estimation, complete genotypes were obtained for 91 of the 100 randomly selected individuals. Of these 91 individuals, four contained allele discrepancies between their two genotypes resulting in seven differing alleles. A single individual accounted for four of the differing alleles suggesting a process error for that individual. The resulting per allele error rate was 0.32% (7/2184).

Pedigree Reconstruction

Sibship reconstruction and parentage assignment analyses performed on the simulated datasets both indicated a high degree of power of the genetic panel to accurately reconstruct full-sibling families and assign parents. For reconstructed full-sibling families composed of at least two individuals, inferred families had a correct partition rate of 91.2% (0.7%) (SE), and assigned parents had an accuracy of 94.2% (0.6%). Accuracies for both methods improved as full-sibling family size increased. For example, full-sibling families composed of at least five individuals resulted in accuracies of 97.7% (0.4%) (sibship), and 96.1% (0.5%) (parentage).

As an additional validation of parentage assignment accuracy, known locations of parents during spawning were compared to natal rivers of assigned families for congruence. Of 101 assigned parents available for capture, 84 were detected during the spawning period that produced their assigned family. Of these 84, 76 were captured in the natal river of the assigned family, resulting in a congruence rate of 90.5%.

Values describing the reconstructed pedigree are shown in Table 4.2. The number of offspring and full-sibling families produced by a spawning year varied considerably (four-fold (offspring) and three-fold (families)). However, the proportions of assigned parents and offspring assigned a parent were relatively constant. This indicates that reproductive processes operating over the three years were relatively consistent.

Environmental Variables

Seasonal flow and temperature varied within each tributary over the three years, but patterns of yearly variation were identical between the two tributaries (Table 4.3). This is not too surprising given the close spatial proximity of the streams (Figure 4.1). Lowest spring and summer flows, and highest summer temperature occurred during 2002. Flows were higher during 2003 and 2004, particularly for summer, and summer temperatures for those years were lower. Xu *et al*. (2010) reported that for this population, spring growth rates increased under warmer temperature-higher flow conditions, while summer growth rates increased under cooler temperature-higher flow conditions. Therefore, for these three years, 2002 had the lowest growth potential, with 2003 and 2004 having relatively equal, higher growth potentials (Table 4.3).

Population Densities and Length Metrics

Population densities varied both within and among tributaries, with YOY values fluctuating more than adult values (Table 4.3). Comparison of YOY densities between the two tributaries revealed that both had low values in 2002 and high values in 2004, but differed in 2003, with the OS having low values and the OL having high values. Xu *et al*. (2010) assessed the effect of density on growth rate in this population, and concluded that increased densities reduced growth rate, with the effect increasing with temperature in both spring and summer. Thus, for these populations, growth rates during 2004 would have experienced the greatest impact from density, particularly during the spring. Age-0

fish inhabiting the OL population should also have experienced density-effects during 2003.

The mean and range of lengths observed during the fall for age-0 and age-1 fish reflected the combined effects of environmental conditions and experienced densities (Figures 2 and 3). The OS tributary, which had the smallest drainage area, was most sensitive to these effects (Figure 4.2). Significant differences in mean size for both age-0 and age-1 fish existed between years ($p < 0.001$ (Age-0), $p = 0.015$ (Age-1)). For both age-classes, differences were generated by the 2002-2003 comparison (post-hoc unequal-N). In 2002, low flow and high summer temperatures resulted in reduced lengths for both age-0 and age-1 individuals despite low densities. Comparatively, in 2003 fish were on average 20 mm (age-0) and 14 mm (age-1) larger. Given that densities for both YOY and adults in those two years were similar (Table 4.3), the difference in length can be almost entirely attributed to improved environmental conditions. The influence of density on length distributions is evident through comparisons of 2003 and 2004 age-0 fish. These two years were more similar environmentally, but differed seven-fold in densities (0.03 (2003) vs. 0.22 (2004). The result was significantly larger fish in 2003 ($p < 0.001$), but a greater range of fish lengths in 2004, with the upper extent of the range being larger than in 2003 (Figure 4.2B). Age-1 fish, which had identical densities in 2003 and 2004, were of nearly identical lengths in the two years.

The OL tributary, which drained an area more than twice that of the OS tributary, displayed similar, but reduced, patterns for age-0 fish (Figure 4.2). However, for age-1 length distributions (Figure 4.3), there were no differences in means between any of the years ($p = 0.587$) despite inferior environmental conditions and equal density in 2002

(Table 4.3). These data imply that the impact of environmental conditions on size-at-age is tempered by drainage area or movement.

Within an age-class, mature fish were always significantly larger than immature fish ($p < 0.015$ for all pair-wise comparisons) (Figure 4.4). Interestingly, means of age-1 immature and mature fish across years were remarkably consistent given the difference in environmental conditions. For the two years when age-0 fish matured, mean lengths were also similar for maturity classes between years, however sample sizes for mature fish were very small (Figure 4.4). While lengths of fish in maturity classes were similar across years, proportions within a class were not. Conditions influencing size-at-age also influenced the proportion of mature fish within an age-class. The relatively poor 2002 year resulted in age-specific mature proportions of 0% (age-0) and 29% (age-1). Comparatively, 2003, the best year in terms of growth opportunity, resulted in mature proportions of 6% (age-0) and 62% (age-1). For 2004, a year environmentally similar to 2003 but with higher densities, mature proportions were 5% (age-0) and 51% (age-1). These data suggest that a minimum threshold must be reached before the decision to mature occurs.

To determine if the threshold was size, growth rate, or condition dependent, size metrics were examined for known mature and immature age-1 fish during the previous spring and summer. Because sampling of the tributaries did not start until summer of 2002, only data for 2003 and 2004 are shown. The two years show exact opposite patterns in the establishment of significant size differences between maturity states (Figure 4.5). In 2003, spring lengths were nearly identical ($p = 0.621$), while spring-to-fall growth rates for fish that would mature were significantly higher ($p = 0.024$). In

2004, spring lengths of future-maturing fish were already significantly greater than lengths of their immature counterparts ($p = 0.002$), while spring-to-fall growth rates did not differ significantly ($p = 0.196$). This implied that the decision to mature was condition dependent, not size dependent, which agrees with the mechanism reported for Atlantic salmon (Rowe *et al.* 1991;Thorpe 1994). Examination of spring and summer condition factors supports this premise, and places the timing of the decision between late March and Early June (Figure 4.6), which is in agreement with the estimate of April reported for Atlantic salmon (Thorpe 1994). Differences between the two maturity classes in condition factors in spring were not significant ($p = 0.297$ (2003), $p = 0.785$ (2004)), while those in summer were either approaching significance ($p = 0.107$ (2003)) or already significant ($p = 0.007$). The lack of significance in 2003 can most likely be attributed to enhanced growth opportunity for all individuals caused by superior environmental conditions and reduced densities. The point of greater importance is that by summer, fish that would mature in the fall were not only longer or getting longer than fish that would remain immature, they were also heavier for a given length.

Survival Analysis

Survival estimates were generated for groups based on maturity status and age. Goodness-of-fit tests indicated that assumptions of the model were not violated for datasets with separate or pooled age-classes evaluated separately for all three years ($p > 0.80$ for all datasets). There was a trend for mature fish to have lower survival than immature fish when assessed over all age-classes (Figure 4.7, panels B, D, and F).

However, survival was only significantly lower for mature fish in 2002. Survival estimates for groups based on maturity status and age resulted in no significant differences among estimates (Figure 4.7, panels A, C, and E). This indicates that reduced survival of mature fish is influenced primarily by age-2 and older fish, for which 90% matured each year on average. Of more importance is the fact that there was no cost of maturation in terms of individual survival for age-0 and age-1 fish.

To determine if there was an indirect cost to early maturation, survival was estimated for offspring using parental length as a covariate. Because survival estimates did not significantly differ among years for age-0 and age-1 fish, data were pooled to increase sample size. Trends in offspring survival indicated a slight negative relationship with both maternal (Figure 4.8A) and paternal (Figure 4.8B) lengths, however, confidence intervals were wide. Thus, for this population there was no indirect cost to maturing at a smaller size in terms of offspring survival.

Age-class contribution

The frequency of parental age-classes contributing to a cohort varied greatly over the three spawning years in both tributaries (Figure 4.9). Family production was skewed towards age-1 individuals during the harsh 2002 year. Comparatively, all four age-classes contributed during the mild 2003 year, with family production relatively equal. In 2004, family production was skewed towards age-2 and age-3 parents, possibly as a result of increased densities experienced by age-0 and age-1 fish. Two other interesting trends were the absence of contribution of age-1 individuals in 2004 after contributing as age-0

individuals in 2003, and the dominance in production of the 2001 cohort in both tributaries over the three spawning years. This dominance becomes even more apparent when contribution is assessed as the number of offspring produced (Figure 4.10).

Discussion

For these headwater brook trout populations, age-at-maturation decreased during years experiencing benign environmental conditions, allowing for increased population recruitment when following harsh years. Abiotic factors were the primary determinant of age-at-maturation, with intraspecific competition having only minimal impact. This was primarily evidenced by the proportion of age-1 fish that matured, which was approximately twice as great during benign years than the harsh year. Comparatively, for the two benign years, the one experiencing higher densities only suffered an 18% reduction in the proportion of mature age-1 fish. This is in agreement with the pattern of increased age-at-maturity during periods of poor growth reported for other salmonid populations (Grover 2005;Morita *et al*. 2005), and supports the premise that density-dependent mechanisms predominate in benign environments whereas density-independent processes predominate in harsh environments (Lobon-Cervia & Rincon 2004).

While the proportion of individuals that matured within an age-class varied with abiotic and biotic conditions, the length at which they did so did not, implying that a threshold must be achieved in order for maturation to occur. Thresholds have been hypothesized to be of two types: physical and overhead (Day & Rowe 2002). Physical

thresholds describe a physical limitation, such as an inability to fit offspring or eggs inside the body, and thus prohibit reproduction from occurring until a critical size is exceeded. Overhead thresholds describe an energetic cost required to become reproductively active. Our data supported the overhead hypothesis given that age-0 fish matured at lengths that were on average 25 mm smaller than immature age-1 fish. Small maturation lengths have also been reported for other brook trout populations (Hutchings 1996;Hutchings 2006), suggesting an overhead threshold is the norm for this species.

Additional support for an overhead threshold was evidenced by mature fish having a higher condition factor in early summer, but not spring. This pattern has been reported for Atlantic salmon as well (Thorpe 1994), and suggests that the "decision" to mature occurs during the spring. However, for the two years analyzed in this study, the timing of increased growth necessary to attain a larger size and higher condition factor was flexible. In 2003, mature fish were of similar length to immature fish in the spring, but had significantly higher spring-to-fall growth rates. In 2004, mature and immature fish had similar growth rates, but mature fish were already significantly larger. This pattern of flexible timing of growth has also been witnessed for age-at-smoltification in a laboratory population of Atlantic salmon (D. Sigourney, unpublished data), and suggests that it may be an adaptation to stochastic environments in which growth opportunity is unpredictable.

An additional mechanism with the potential to affect threshold exceedence, but not considered in this study, is the existence of substantial size distributions for age-0 fish generated by differences in emergence timing (Hutchings 1996). Emergence date is determined by parental spawning date, temperature conditions within the redd, and

individual behavior (Curry *et al*. 1991;Curry *et al.*1995). The combined effect of these factors results in emergence occurring over an extended period of time (> 2 months) (Snucins *et al.*1992). Furthermore, individual ranks within a length distribution over the lifespan of a cohort are largely maintained (B. Letcher, unpublished data). Because early emerging fish have a competitive size advantage (Einum & Nislow 2005), they may be more likely to exceed the maturation threshold sooner, and thus have a decreased age-at-maturity.

Earlier age-at-maturity is predicted to occur when benefits of early maturation exceed costs. For these brook trout populations, the primary factor influencing the decision to mature appeared to be the benefits gained by increased size (greater fecundity for females, increased mate acquisition for males). Support for this inference comes from the fact that fecundity predicted using mean length of immature and mature fish within an age-class would result in mature fish producing an average of 38% (age-0) and 33% (age-1) more eggs than immature fish. Furthermore, our prediction of increased mortality costs for both early-maturing individuals and their offspring was not supported. In fact, survival estimates for mature fish in this study are in direct contrast to those reported for three Newfoundland brook trout populations, for which mature fish were reported to have significantly higher mortality rates than immature fish (Hutchings 1994;Hutchings *et al*. 1999). Survival differences between these populations can most likely be attributed to the increased length and harshness of the winter season and the reduced growth opportunity present in those populations, which resided at the northern edge of the species' range. Thus, costs and benefits differed between these populations, as evidenced by the

increased optimal age-at-maturation for the Newfoundland populations (>3 years) (Hutchings 1996).

The effect of parental length on offspring survival also failed to incur a cost, again contradicting our prediction. The lack of an effect from paternal lengths was not too surprising, as data testing the so called "good genes" hypothesis has reported the same result for both brown trout (Jacob *et al*. 2007) and Atlantic salmon (Garant *et al*. 2002). However, the lack of a maternal length effect was surprising as egg size has been linked to enhanced survival (Einum & Fleming 1999;Einum & Fleming 2000a), and is positively correlated with female length (Morita 1998). Female brook trout inhabiting this population also exhibited a positive correlation between female size and egg diameter ($Y = 2.2274+0.0109*X$, $N = 42$, $r^2 = 0.39$, $p < 0.001$ (J. Coombs, unpublished data)), however, the relationship between female size and egg dry weight resulted in no correlation ($Y = 6.507 + 0.0366*X$, $N = 10$, $r^2 = 0.17$, $p = 0.232$ (H. Wang, unpublished data)). This supports the absence of a maternal size effect on offspring survival, and suggests that differences in egg size seen in this population may be attributed to increased water weight of larger female eggs as an artifact of sampling on a single day given that spawning progresses temporally from larger to smaller sized females (Blanchfield & Ridgway 2005).

One cost of maturation that did occur was the failure of mature individuals to spawn successfully in multiple years. For the three cohorts assessed in this study, only 4 out of 78 assigned parents were successful repeat spawners (3 females, 1 male) (J. Coombs, unpublished data). The reason was not due to failure to mature again (18 out of 19 fish were observed in a mature state in two consecutive years (J. Coombs, unpublished

data)), thus implying that repeat spawners were either less competitive (males), or their offspring were less viable (females). Both of these reasons could be side effects of lipid loss sustained from the first maturation (Hutchings *et al.*1999), and an inability to recover these losses prior to the subsequent spawning year (Berg *et al.*1998). Regardless of the reason, this indicates that for this population, fitness at first spawning is approximately equivalent to lifetime reproductive success. Predicted fecundities based on mean length of mature fish within an age-class were 27 (age-0), 68 (age-1), and 130 (age-2). The fact that individuals still matured early suggests that overall survival rates are low and unpredictable, and that early maturation must have fitness advantages. For this study, this was most apparent in the benign, low density year during which age-0 individuals contributed to 5 out of 33 families, accounting for 9% of the number of offspring.

The ability of stream salmonid populations to recover from disturbance events through increased recruitment has been hypothesized to be a mechanism to increase probabilities of population persistence (Vincenzi *et al*. 2008). For the brook trout populations in this study, the occurrence of an environmentally benign year subsequent to a harsh one increased individual growth opportunity through both density-independent and density-dependent processes. This in turn resulted in decreased age-at-maturation, and reproductive contribution coming from all age-classes, thus increasing recruitment. Furthermore, effective population sizes (N_e) increased during the benign years (Coombs 2010), most-likely as a result of added contribution from younger age-classes acting as a form of genetic compensation (Palstra & Ruzzante 2008). The merits of increased N_e have been well documented for population viability (Reed 2005;Saccheri *et al*. 1998) and adaptive potential (Frankham *et al*. 1999). Furthermore, the harsh year produced a

"cohort effect" (Lindstrom & Kokko 2002) for the 2001 year class, resulting in that cohort accounting for the overwhelming majority of offspring produced over the three study years. Lindstrom and Kokko (2002) reported that cohort effects act to stabilize populations with non-linear density-dependencies, relatively high potential growth rates, and overlapping generations, all of which exist for these brook trout populations.

These data support the hypothesis that reduced age-at-maturity can buffer against disturbance events through increased recruitment potential, and thus act as a mechanism to increase population persistence. However, the ability of this mechanism to function is dependent upon the occurrence of benign years following harsh years. Given that climate change is predicted to substantially alter hydrologic regimes within headwater streams across the native range of brook trout (Marshall & Randhir 2008), population persistence is likely to be negatively impacted if environmental conditions constrain the effectiveness of this mechanism. Future research is needed in this area to determine how climate change will impact population persistence, and particularly this mechanism.

In summary, we found support for the hypothesis that reduced age-at-maturity acts as a mechanism to facilitate population recovery through increased recruitment. The primary determinant of age-at-maturity appeared to be fitness benefits achieved from increased size-at-age, as there were no survival costs to either the maturing individual or their offspring. The only cost detected for these populations was the lack of success for repeat spawners, resulting in the population displaying a predominately semelparous reproductive strategy. In spite of decreased size-dependent benefits, selection still appeared to favor earlier maturation. This most likely reflects the unpredictability in survival for these populations as a function of varying environmental conditions. The

ability of this mechanism to function is constrained by individual growth opportunities also determined by environmental conditions. Given headwater stream hydrologic regimes are predicted to be altered by climate change, the effectiveness of this mechanism to increase population persistence may be compromised. More research is required in this area to assess how such changes will manifest.

Table 4.1. Single locus summary statistics for brook trout cohorts inhabiting the open-small (OS) and open-large (OL) populations. Measures for each locus are: (A_O) observed number of alleles; (H_O) observed heterozygosity; (H_E) expected heterozygosity; (f_{is}) an analogue of Wright's F_{IS} statistic; (p) probability of departure from Hardy-Weinberg expectations in the direction of heterozygote deficiency. Bold p values indicate significant genotypic departures from expected Hardy-Weinberg equilibrium when evaluated using a sequential Bonferroni correction for multiple tests (k = 48, α = 0.05).

Population	Cohort		SfoC24	SfoC88	SfoC113	SfoC115	SfoC129	SfoD75	SfoD100
OS	2002	A_O	4	3	7	4	4	7	5
	(n=54)	H_O	0.26	0.43	0.96	0.46	0.41	0.69	0.46
		H_E	0.28	0.36	0.79	0.52	0.40	0.71	0.57
		f_{is}	0.08	-0.18	-0.22	0.12	-0.02	0.04	0.19
		p	0.49	1.00	1.00	0.03	0.49	0.40	0.15
	2003	A_O	3	4	5	4	5	7	4
	(n=40)	H_O	0.78	0.45	0.53	0.90	0.88	0.85	0.85
		H_E	0.66	0.53	0.49	0.52	0.81	0.71	0.62
		f_{is}	-0.17	0.14	-0.06	-0.74	-0.08	-0.20	-0.39
		p	0.96	0.17	0.40	1.00	0.80	0.99	1.00
	2004	A_O	4	4	8	6	5	7	5
	(n=201)	H_O	0.35	0.39	0.82	0.77	0.72	0.87	0.45
		H_E	0.31	0.39	0.79	0.70	0.66	0.75	0.58
		f_{is}	-0.14	0.00	-0.04	-0.09	-0.09	-0.16	0.22
		p	1.00	0.98	0.87	1.00	0.97	1.00	0.08
	2005	A_O	4	3	7	6	5	8	5
	(n=73)	H_O	0.34	0.66	0.78	0.64	0.51	0.73	0.77
		H_E	0.32	0.50	0.78	0.56	0.51	0.67	0.65
		f_{is}	-0.06	-0.32	0.00	-0.15	0.00	-0.08	-0.18
		p	0.70	1.00	0.64	0.99	0.67	0.88	1.00

Table 4.1. (continued)

Population	Cohort		SsaD237	SfoB52	SfoC38	SfoC86	SfoD91a	Average
OS	2002	A_O	9	3	3	2	4	4.58
	(n=54)	H_O	0.87	0.65	0.20	0.41	0.81	0.55
		H_E	0.74	0.58	0.22	0.47	0.59	0.52
		f_{is}	-0.18	-0.12	0.06	0.14	-0.39	-0.04
		p	1.00	0.97	0.45	0.24	1.00	
	2003	A_O	9	3	3	2	5	4.5
	(n=40)	H_O	0.83	0.63	0.45	0.45	0.65	0.69
		H_E	0.75	0.66	0.39	0.38	0.55	0.59
		f_{is}	-0.10	0.05	-0.17	-0.19	-0.19	-0.17
		p	0.60	0.36	0.93	0.96	0.96	
	2004	A_O	13	4	3	4	7	5.83
	(n=201)	H_O	0.91	0.74	0.40	0.46	0.71	0.63
		H_E	0.84	0.61	0.46	0.40	0.65	0.59
		f_{is}	-0.08	-0.21	0.13	-0.15	-0.10	-0.06
		p	0.99	1.00	0.17	0.99	0.29	
	2005	A_O	10	5	3	4	7	5.58
	(n=73)	H_O	0.92	0.52	0.25	0.53	0.67	0.61
		H_E	0.84	0.50	0.22	0.45	0.67	0.56
		f_{is}	-0.09	-0.05	-0.10	-0.19	0.00	-0.10
		p	0.43	0.79	1.00	0.98	0.66	

Table 4.1. (continued)

Population	Cohort		SfoC24	SfoC88	SfoC113	SfoC115	SfoC129	SfoD75	SfoD100
OL	2002	A_O	4	5	8	9	5	9	6
	(n=161)	H_O	0.13	0.60	0.71	0.36	0.55	0.74	0.62
		H_E	0.12	0.58	0.74	0.37	0.50	0.78	0.64
		f_{is}	-0.05	-0.04	0.03	0.03	-0.11	0.05	0.03
		p	1.00	0.86	0.27	0.01	0.94	0.01	0.58
	2003	A_O	4	5	8	10	5	11	6
	(n=208)	H_O	0.15	0.56	0.81	0.34	0.46	0.75	0.63
		H_E	0.16	0.54	0.77	0.38	0.47	0.78	0.65
		f_{is}	0.06	-0.03	-0.05	0.13	0.03	0.04	0.04
		p	0.24	0.88	0.96	0.26	0.34	0.03	0.57
	2004	A_O	4	4	9	9	5	10	8
	(n=185)	H_O	0.08	0.51	0.71	0.22	0.48	0.77	0.62
		H_E	0.08	0.52	0.75	0.25	0.46	0.80	0.61
		f_{is}	-0.03	0.02	0.05	0.13	-0.04	0.05	-0.01
		p	1.00	0.56	0.19	0.53	0.17	0.18	0.15
	2005	A_O	3	4	7	7	5	7	5
	(n=51)	H_O	0.16	0.55	0.67	0.22	0.45	0.78	0.61
		H_E	0.15	0.56	0.76	0.25	0.48	0.80	0.58
		f_{is}	-0.06	0.02	0.12	0.15	0.05	0.02	-0.04
		p	1.00	0.47	0.13	0.08	0.10	0.49	0.60

Table 4.1. (continued)

Population	Cohort		SsaD237	SfoB52	SfoC38	SfoC86	SfoD91a	Average
OL	2002	A_O	18	6	3	6	9	7.33
	(n=161)	H_O	0.76	0.68	0.61	0.39	0.73	0.57
		H_E	0.80	0.65	0.51	0.41	0.77	0.57
		f_{is}	0.05	-0.04	-0.19	0.05	0.05	-0.01
		p	**0.00**	0.84	0.96	0.46	0.17	
	2003	A_O	24	7	3	5	10	8.17
	(n=208)	H_O	0.88	0.66	0.61	0.51	0.65	0.58
		H_E	0.89	0.71	0.53	0.50	0.75	0.60
		f_{is}	0.01	0.06	-0.14	-0.03	0.13	0.02
		p	0.63	0.30	1.00	0.79	0.01	
	2004	A_O	22	7	3	5	10	8
	(n=185)	H_O	0.81	0.72	0.36	0.52	0.72	0.54
		H_E	0.85	0.74	0.40	0.50	0.74	0.56
		f_{is}	0.04	0.02	0.10	-0.05	0.02	0.02
		p	0.08	0.52	0.08	0.77	0.05	
	2005	A_O	16	5	3	5	6	6.08
	(n=51)	H_O	0.71	0.71	0.49	0.57	0.76	0.56
		H_E	0.74	0.71	0.45	0.52	0.72	0.56
		f_{is}	0.05	0.01	-0.09	-0.10	-0.06	0.01
		p	0.44	0.57	0.84	0.85	0.82	

Table 4.2. Descriptive measures for reconstructed full-sibling families containing a minimum of two individuals for brook trout populations inhabiting the open-small and open-large tributaries. SY = Spawning year, N_O = Number of offspring, N_F = Number of full-sibling families, N_P = Number of assigned parents, AP = Proportion of assigned parents, AO = Proportion of offspring assigned a parent.

SY	$\mathbf{N_O}$	$\mathbf{N_F}$	$\mathbf{N_P}$	**AP**	**AO**
2002	298	71	38	0.27	0.61
2003	401	84	44	0.26	0.60
2004	112	28	21	0.38	0.65

Table 4.3. Average daily discharge and temperature over the spring and summer seasons, and number and density of brook trout captured during the fall sample in the open-small (OS) and open-large (OL) tributaries.

		Discharge (m^3/s)		Temperature (^oC)		Count			Density	
Stream	**Year**	**Spring[1]**	**Summer[2]**	**Spring[1]**	**Summer[2]**	**YOY**	**Adults**	**Area[3]**	**YOY**	**Adults**
OS	2002	0.015	0.002	7.9	15.9	25	12	613	0.04	0.02
	2003	0.031	0.011	7.6	15.4	24	41	818	0.03	0.05
	2004	0.025	0.008	8.4	15.0	168	35	770	0.22	0.05
OL	2002	0.042	0.011	7.5	15.1	50	82	824	0.06	0.10
	2003	0.066	0.030	6.7	14.5	138	110	949	0.15	0.12
	2004	0.054	0.025	7.6	14.4	156	86	926	0.17	0.09

[1]March 1st to June 22nd

[2]June 23rd to October 21st

[3]Calculated using wetted widths derived from average daily summer discharge value

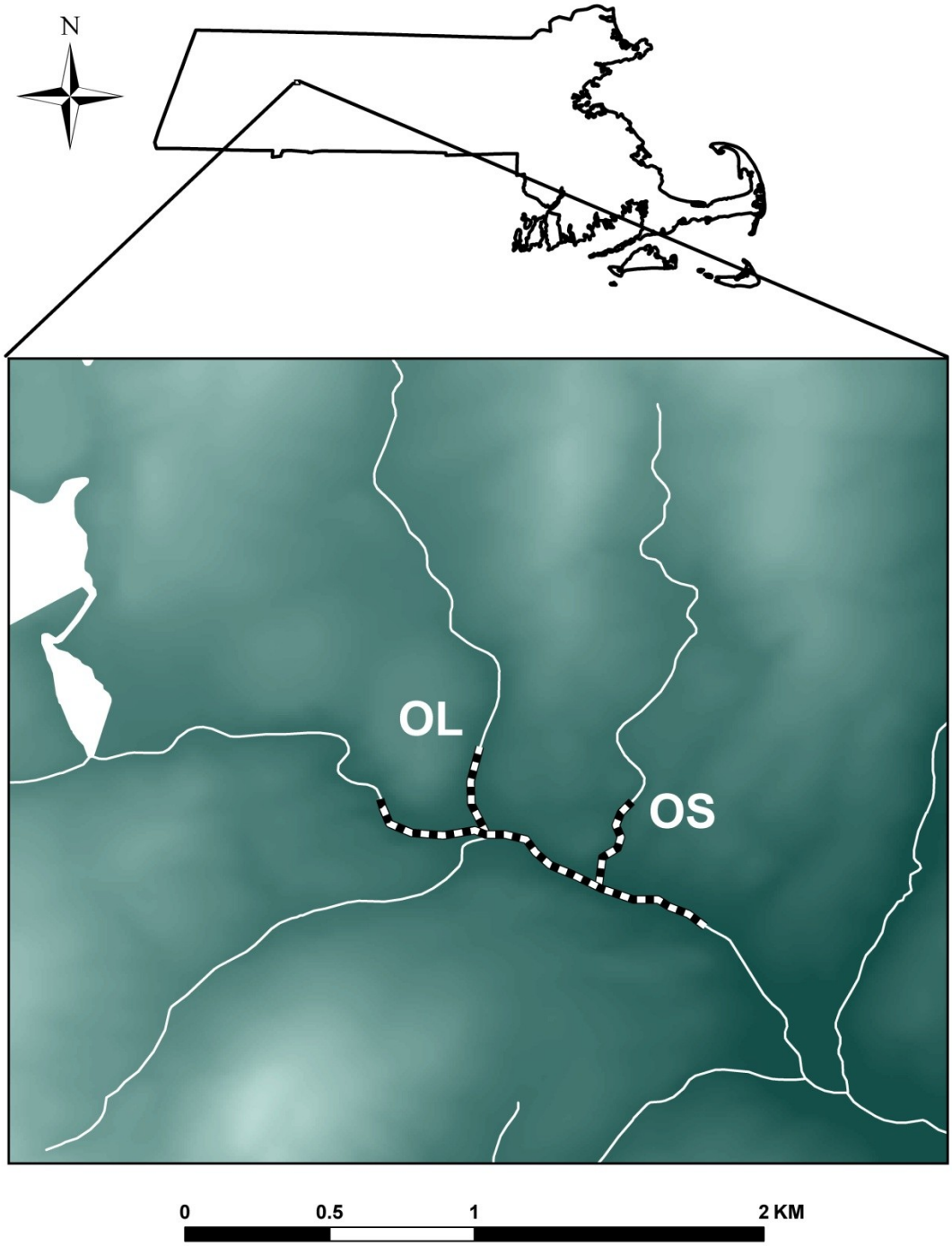


Figure 4.1. Map showing the West Brook drainage located in Whately, MA, USA. OL = Open-Large tributary; OS = Open-Small tributary. Dashed line indicates sampled reaches.

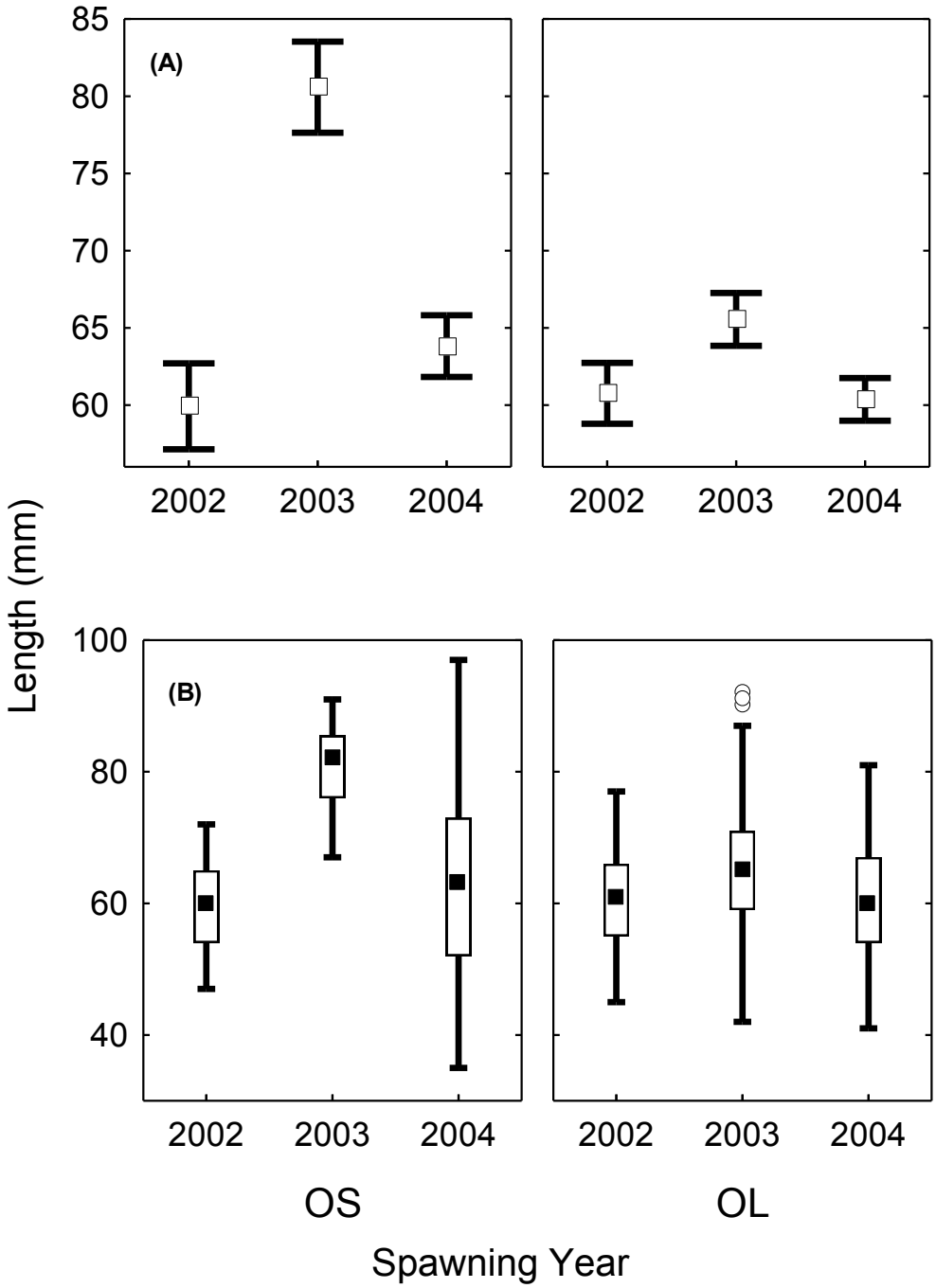


Figure 4.2. Sizes of young-of-year (YOY) brook trout caught in the fall sample of 2002, 2003, and 2004 in the open-small (OS) and open-large (OL) tributaries. (A) Mean and 95% confidence intervals. (B) Box-plot showing median (square), middle 50% (box), non-outlier range (lines), and outliers (circles). OS sample sizes: 25 (2002), 24 (2003), 168 (2004); OL sample sizes: 50 (2002), 138 (2003), 156 (2004).

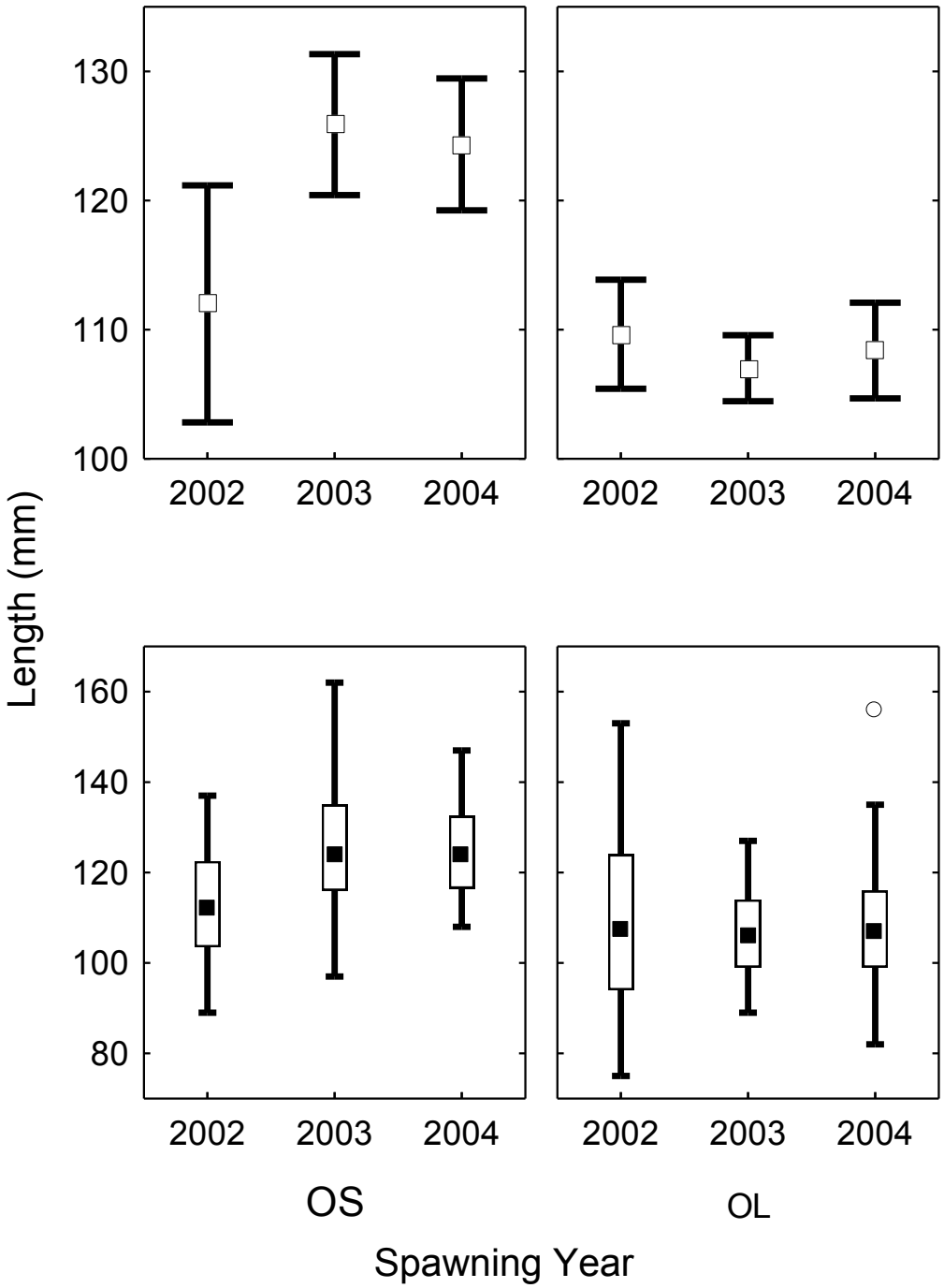


Figure 4.3. Sizes of age-1 brook trout caught in the fall sample of 2002, 2003, and 2004 in the open-small (OS) and open-large (OL) tributaries. (A) Mean and 95% confidence intervals. (B) Box-plot showing median (square), middle 50% (box), non-outlier range (lines), and outliers (circles). OS sample sizes: 12 (2002), 33 (2003), 20 (2004); OL sample sizes: 73 (2002), 60 (2003), 55 (2004).

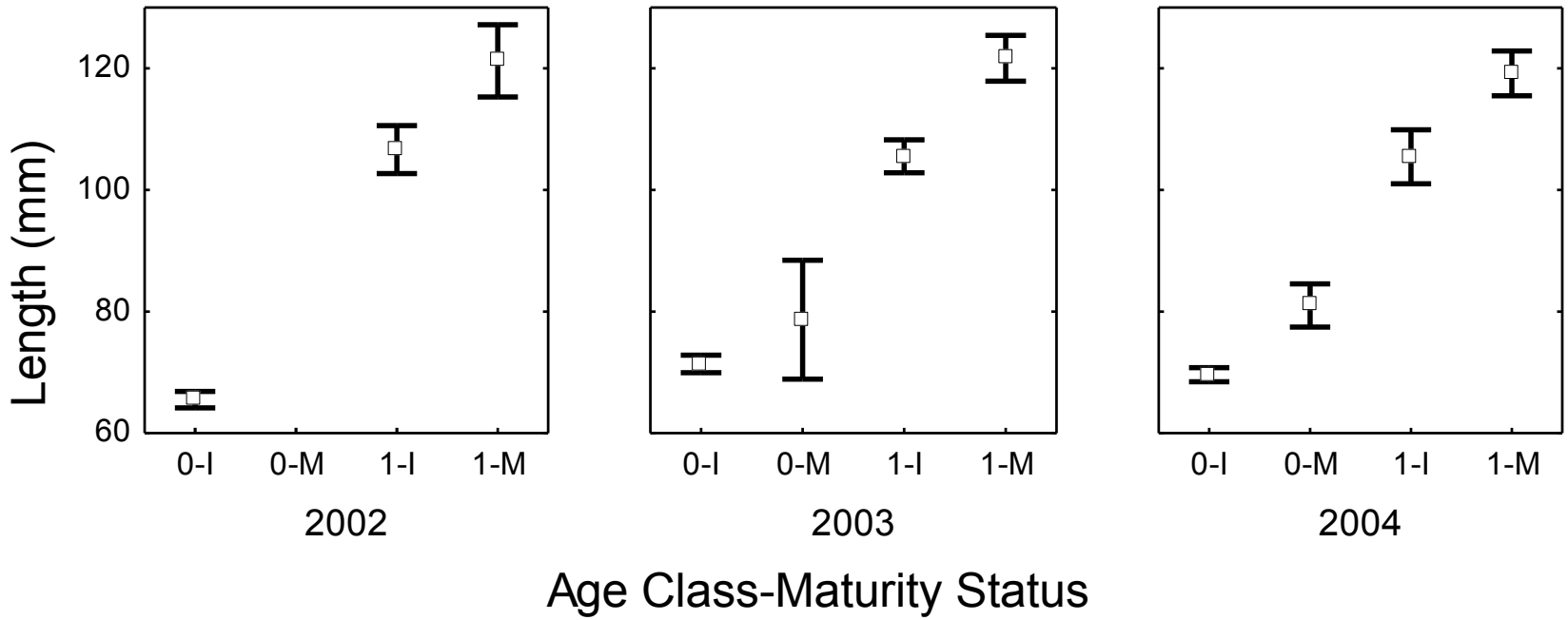


Figure 4.4. Mean size of immature (I) and mature (M) age-0 and age-1 brook trout inhabiting the open-small and open-large tributaries during the 2002, 2003, and 2004 spawning seasons. Sample sizes for 0-I, 0-M, 1-I, 1-M are 41, 0, 66, 27 (2002); 120, 8, 38, 63 (2003); and 171, 9, 37, 38 (2004).

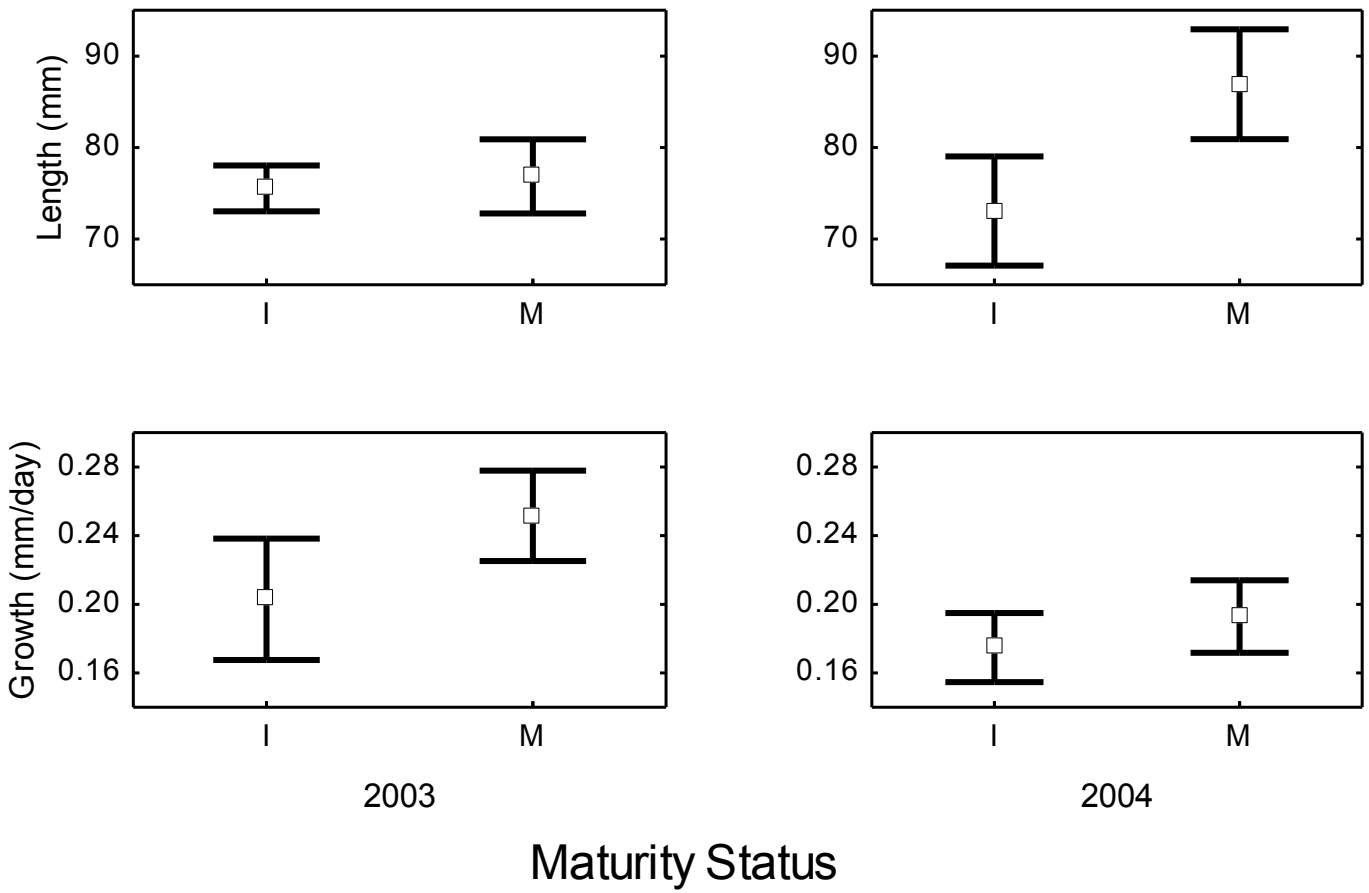


Figure 4.5. Length in spring, and spring-to-fall growth rate of immature (I) and mature (M) age-1 brook trout during the 2003 and 2004 spawning years. Sample sizes: 13 (2003-I), 21 (2003-M), 16 (2004-I), 13 (2004-M).

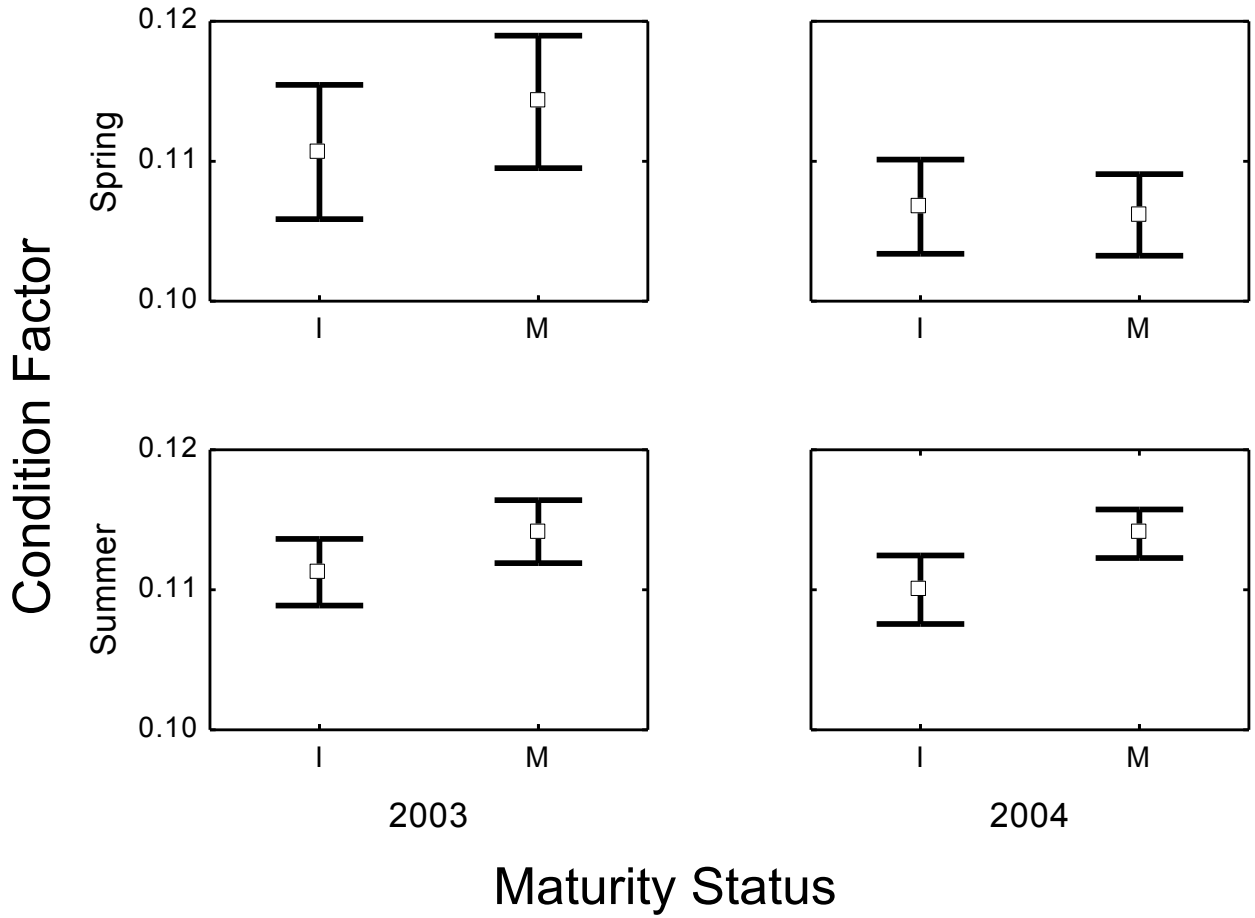


Figure 4.6. Condition factor in the spring and summer for age-1 immature (I) and mature (M) brook trout during the 2003 and 2004 spawning years. Spring sample sizes: 13 (2003-I), 21 (2003-M), 16 (2004-I), 13 (2004-M); Summer sample sizes: 22 (2003-I), 44 (2003-M), 26 (2004-I), 36 (2004-M).

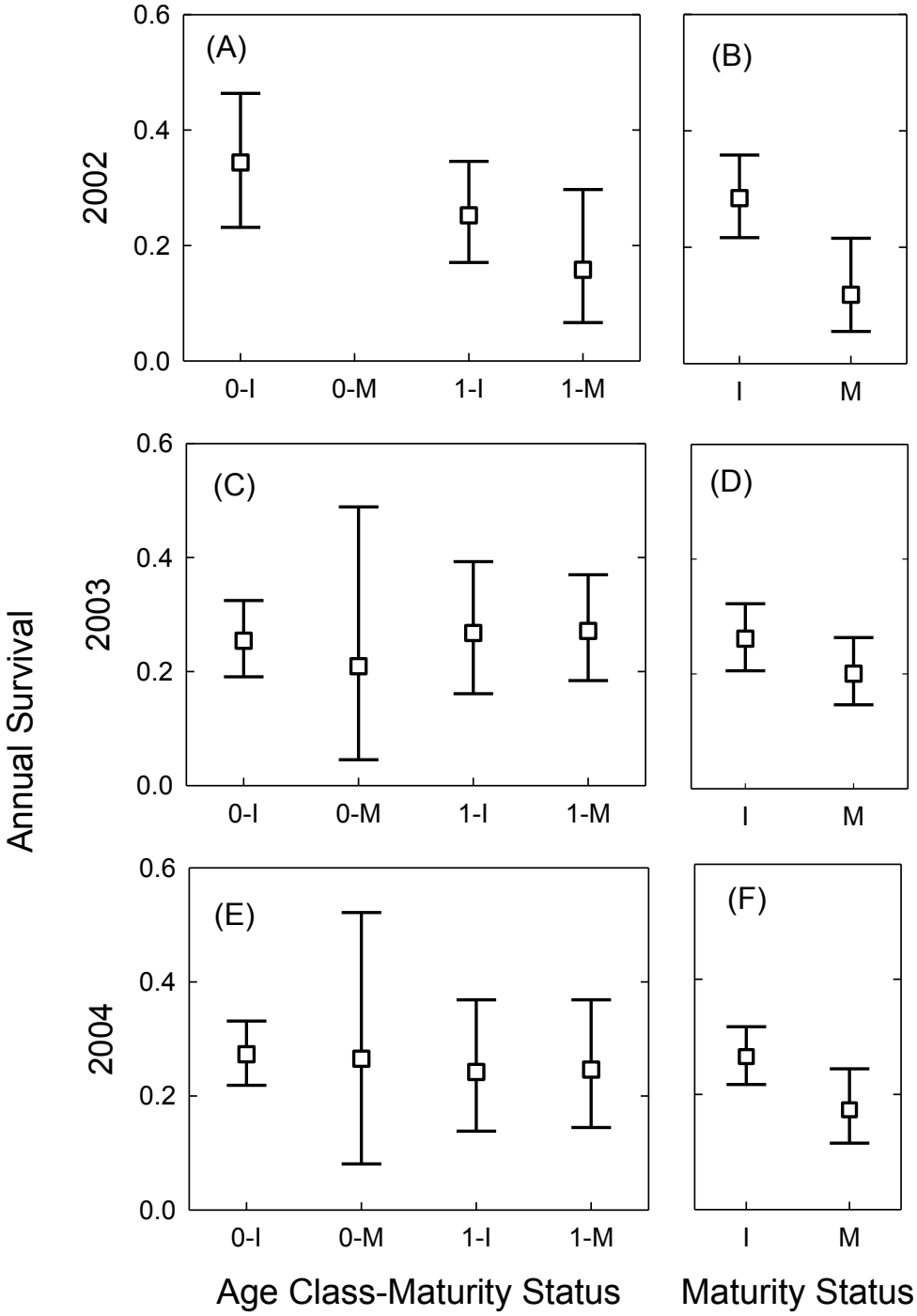


Figure 4.7. Annual survival estimates (mean and 95% CI's) for immature (I) and mature (M) brook trout detected during the 2002, 2003, and 2004 spawning periods. Graphs A, C, and E show survival for age-0 and age-1 fish; Graphs B, D, and F show survival for all age-classes pooled. Sample sizes for 0-I, 0-M, 1-I, 1-M were: 41, 0, 66, 27 (2002); 120, 8, 38, 63 (2003); 171, 9, 37, 38 (2004). Sample sizes for pooled I and M were: 109 and 45 (2002); 160 and 137 (2003); 215 and 97 (2004).

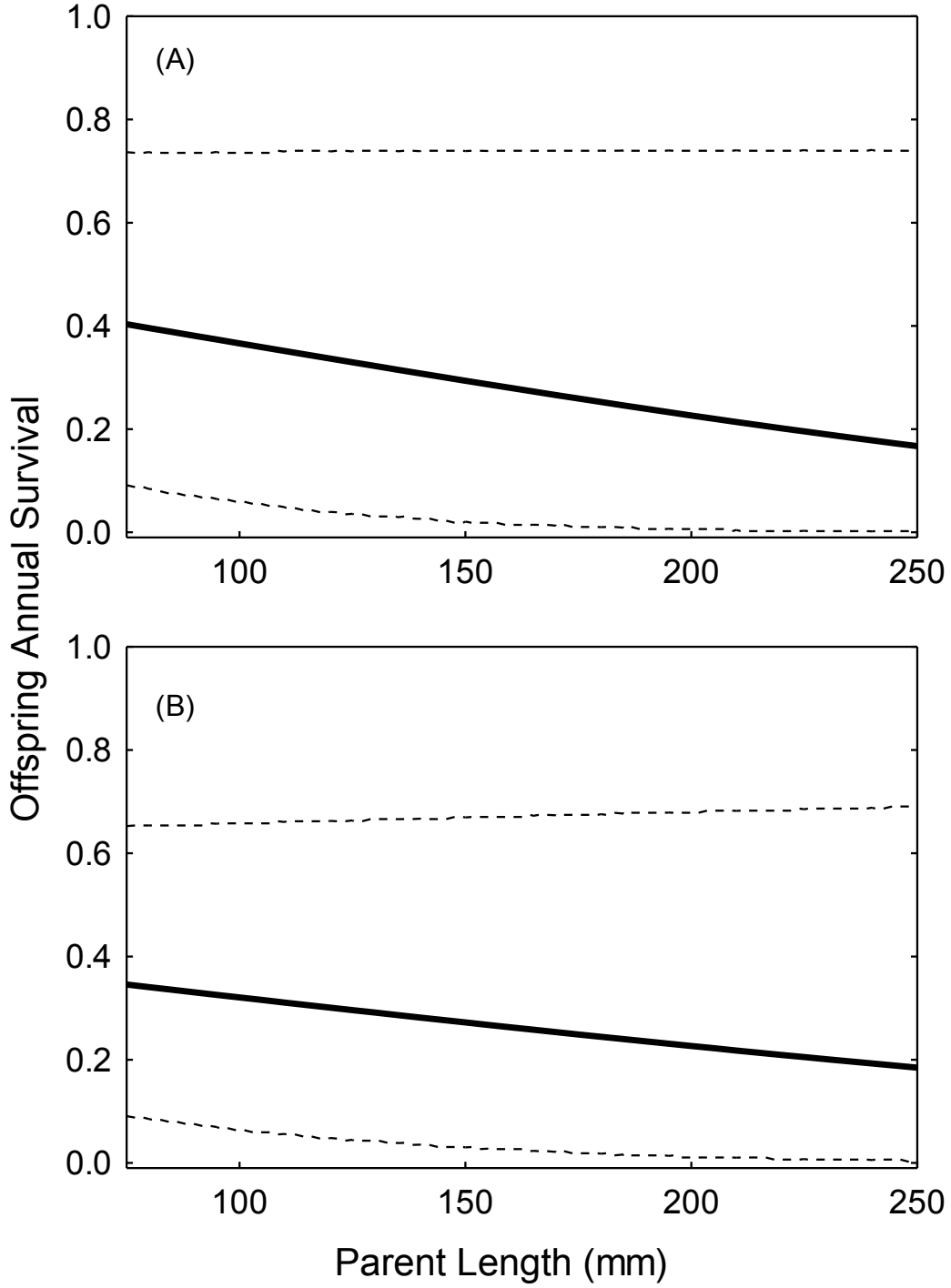


Figure 4.8. Effect of maternal (A) and paternal (B) lengths on offspring survival. Solid line represents offspring survival estimate; dashed lines represent 95% confidence intervals. (A) Y = -0.0009*X+0.8692, based on 600 observations of 256 offspring; (B) Y = -0.0006*X+0.8166, based on 663 observations of 290 offspring; Both had 17 sampling occasions.

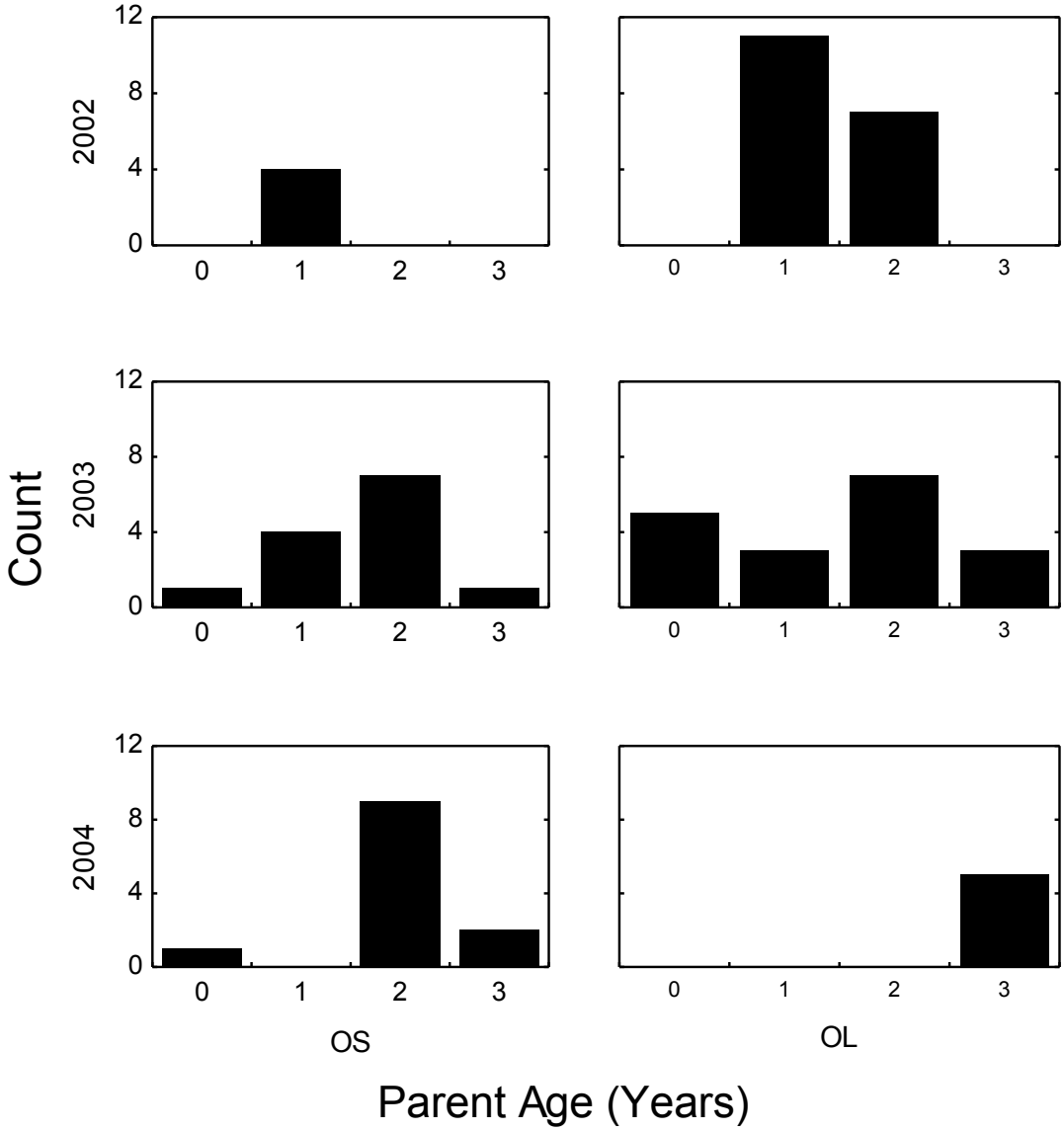


Figure 4.9. Ages of contributing parents for the 2002-2004 brook trout cohorts inhabiting the open-small (OS) and open-large (OL) tributaries.

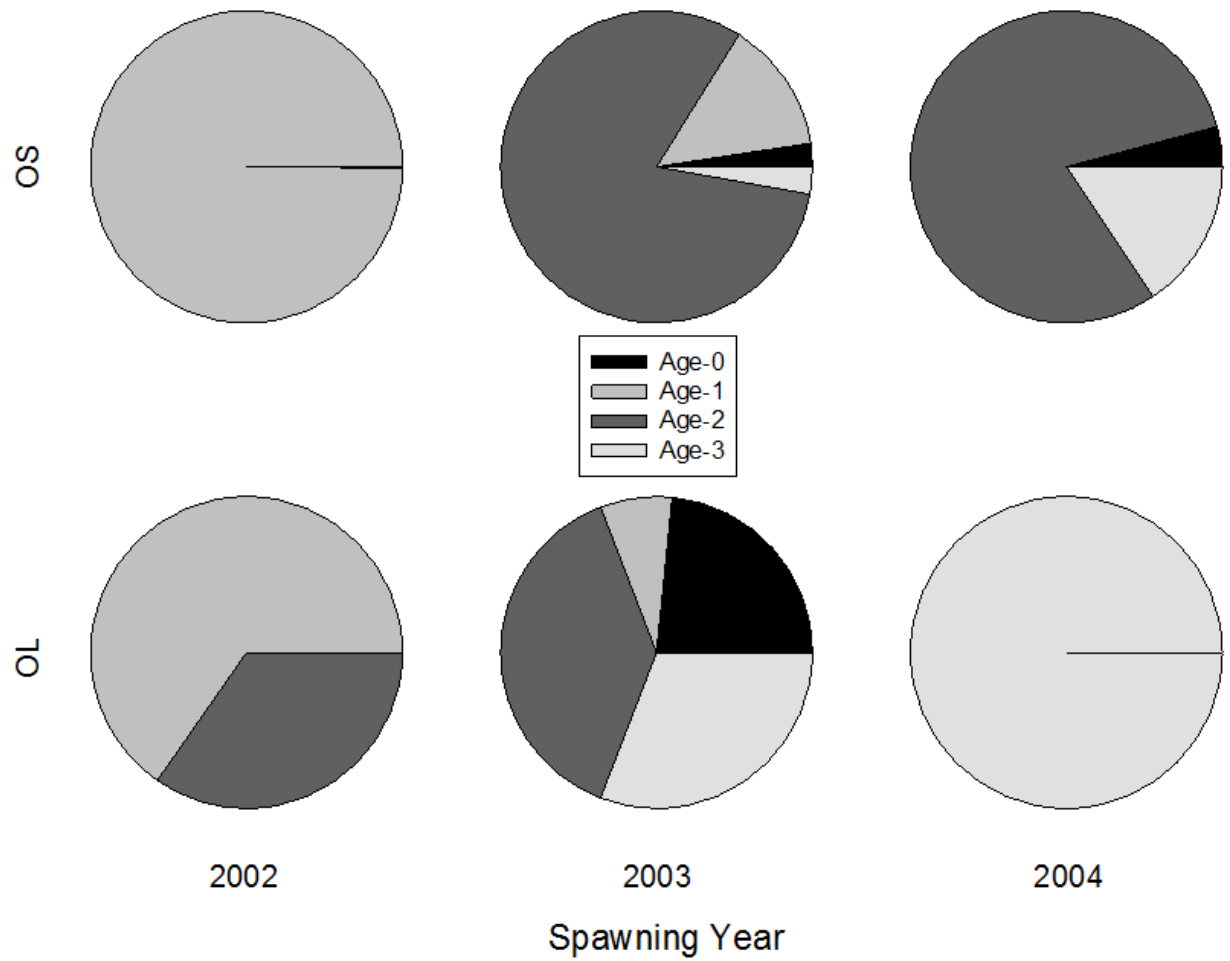


Figure 4.10. Proportion of offspring contributed by parental age-class for the 2002, 2003, and 2004 spawning years in the open-small (OS) and open-large (OL) brook trout populations.

CHAPTER V

SYNOPSIS

Small populations are predicted to suffer increased extinction probability as a consequence of effects suffered from reduced genetic diversity, and demographic and environmental stochasticity. However, many small populations continue to persist, suggesting adaptations may have evolved that in turn gave rise to mechanisms acting to counter these effects. Such mechanisms would likely involve reproduction, as it directly influences genetic diversity and census size of a population. Genetic diversity is conserved when a large proportion of the population contributes reproductively, which is ultimately determined by patterns in mate choice. Census size is increased through recruitment, which is determined by the number of reproducing individuals, which is itself a function of age-at-maturity. Thus in small, persistent populations, mate choice patterns would be predicted to result in a large-proportion of adults contributing reproductively, and age-at-maturation would be predicted to be flexible to account for environmental stochasticity.

To assess these predictions, reproductive patterns were determined for two wild brook trout (*Salvelinus fontinalis*) populations inhabiting headwater streams. For these populations, mate choice was size-assortative, with males and females within a pair having approximately equal length. This pattern most-likely resulted from males selecting larger females to benefit from their increased fecundity, and females selecting larger males to benefit from their ability to deter egg cannibalism. The result of this mate

choice pattern was a large proportion of individuals (0.8) mating only one time during a reproductive cycle. This parity in reproductive contribution produced a mean effective population size (N_e) to census population size (N_c) ratio of 0.49, a value four to five times larger than mean ratios reported for 165 (0.14) and 102 (0.10) different species. These data suggest that size-assortative mate choice patterns also produce a mechanism that acts to conserve genetic diversity.

For age-at-maturity, selection appeared to favor early maturation, most-likely as a result of high and unpredictable mortality rates. Benefits of early maturation were based on body size and its positive correlation with fecundity. Larger females had greater fecundity which directly increased their fitness, while larger males mated with larger females (size-assortative mating) which indirectly increased their fitness. For early maturing fish, there was no evidence for direct costs in terms of survival compared to their immature counterparts, or indirect costs in terms of their offspring's survival compared to those produced by larger parents. One apparent cost that did manifest was a lack of successful spawning in multiple years (5%), essentially rendering these brook trout semelparous.

The age at which a brook trout first matured in these populations ranged from zero to two years, and was primarily determined by growth opportunity mediated through environmental conditions. Maturation appeared to be dependent upon surpassing an energetic threshold value, as mature age-0 fish occurred at a smaller length than immature age-1 fish. Additionally, mature fish within an age-class were significantly longer and heavier than their immature counterparts. The ability of a fish to surpass this threshold was flexible, with different means employed in different years, suggesting an

adaptation to capitalize on growth opportunities in an unpredictable environment. Harsh environmental conditions resulted in reduced growth opportunity, and delayed age-at-maturation. Benign environmental conditions resulted in increased growth opportunity, and earlier age-at-maturation. This was particularly apparent for a benign year following a harsh year, as decreased densities further enhanced growth opportunity. These data suggest that flexible age-at-maturation also results in a mechanism that acts to increase population recruitment after the occurrence of a stochastic event.

Adaptations for mate-choice and flexible age-at-maturity appear to have evolved in these headwater brook trout populations. These adaptations in turn gave rise to mechanisms acting to increase a population's persistence probability through conservation of genetic diversity and increased recruitment potential. Ultimately, both mechanisms are dependent upon individual growth patterns. Size-assortative mating requires that a range of individual lengths be present in the population during reproduction. Flexible age-at-maturity requires that an energetic threshold be surpassed in order for an individual to mature. Given this, the efficacy of both mechanisms is ultimately linked to growth opportunity mediated through environmental conditions. Thus, changes in headwater habitat conditions predicted to occur as a result of climate change could compromise these mechanisms and render brook trout populations more susceptible to local extirpation.

APPENDIX

ALLELE FREQUENCIES FOR CHAPTER II SIMULATIONS

Locus	Allele	Frequency	Locus	Allele	Frequency
C113-Allele 1	125	0.00046	D100-Allele 1	206	0.02493
C113-Allele 1	128	0.04620	D100-Allele 1	211	0.15897
C113-Allele 1	132	0.01532	D100-Allele 1	215	0.19030
C113-Allele 1	135	0.31770	D100-Allele 1	219	0.20608
C113-Allele 1	138	0.27219	D100-Allele 1	224	0.07708
C113-Allele 1	142	0.09584	D100-Allele 1	228	0.01761
C113-Allele 1	145	0.09584	D100-Allele 1	233	0.17887
C113-Allele 1	148	0.04643	D100-Allele 1	237	0.00618
C113-Allele 1	152	0.03019	D100-Allele 1	241	0.06404
C113-Allele 1	158	0.04460	D100-Allele 1	249	0.03957
C113-Allele 1	162	0.03522	D100-Allele 1	253	0.03568
D75-Allele 1	176	0.05169	D100-Allele 1	257	0.00069
D75-Allele 1	181	0.00709	C115-Allele 1	232	0.05764
D75-Allele 1	185	0.00091	C115-Allele 1	238	0.03225
D75-Allele 1	189	0.00023	C115-Allele 1	242	0.00503
D75-Allele 1	193	0.10567	C115-Allele 1	244	0.00023
D75-Allele 1	197	0.01350	C115-Allele 1	246	0.00549
D75-Allele 1	201	0.03866	C115-Allele 1	258	0.00160
D75-Allele 1	206	0.05993	C115-Allele 1	302	0.00114
D75-Allele 1	210	0.13747	C115-Allele 1	306	0.00549
D75-Allele 1	214	0.19236	C115-Allele 1	310	0.01647
D75-Allele 1	218	0.05375	C115-Allele 1	314	0.00663
D75-Allele 1	222	0.28088	C115-Allele 1	322	0.00069
D75-Allele 1	226	0.05787	C115-Allele 1	326	0.02424
C88-Allele 1	177	0.04552	C115-Allele 1	328	0.00046
C88-Allele 1	180	0.00091	C115-Allele 1	330	0.04140
C88-Allele 1	183	0.35522	C115-Allele 1	334	0.27928
C88-Allele 1	186	0.24177	C115-Allele 1	338	0.12626
C88-Allele 1	189	0.03500	C115-Allele 1	342	0.13769
C88-Allele 1	192	0.09790	C115-Allele 1	344	0.04231
C88-Allele 1	195	0.21706	C115-Allele 1	346	0.12214
C88-Allele 1	201	0.00663	C115-Allele 1	350	0.09081
			C115-Allele 1	354	0.00274

APPENDIX (continued)

Locus	Allele	Frequency	Locus	Allele	Frequency
C129-Allele 1	223	0.11963	D237-Allele 1	454	0.07236
C129-Allele 1	229	0.14021	D237-Allele 1	466	0.03267
C129-Allele 1	232	0.22461	D237-Allele 1	470	0.00097
C129-Allele 1	236	0.51441	Locus-9	200	0.05935
C129-Allele 1	239	0.00114	Locus-9	204	0.00078
C24-Allele 1	110	0.04209	Locus-9	208	0.04731
C24-Allele 1	113	0.23994	Locus-9	212	0.18812
C24-Allele 1	116	0.43115	Locus-9	216	0.00728
C24-Allele 1	119	0.06976	Locus-9	220	0.16035
C24-Allele 1	122	0.05581	Locus-9	224	0.11765
C24-Allele 1	170	0.16125	Locus-9	228	0.28305
D237-Allele 1	276	0.01162	Locus-9	232	0.12060
D237-Allele 1	280	0.01839	Locus-9	236	0.01551
D237-Allele 1	284	0.00823	Locus-10	210	0.07428
D237-Allele 1	288	0.00169	Locus-10	214	0.13962
D237-Allele 1	292	0.03437	Locus-10	218	0.26252
D237-Allele 1	296	0.00024	Locus-10	222	0.07446
D237-Allele 1	300	0.00871	Locus-10	226	0.04958
D237-Allele 1	304	0.00048	Locus-10	230	0.29601
D237-Allele 1	308	0.04066	Locus-10	234	0.10353
D237-Allele 1	373	0.00024	Locus-11	220	0.22123
D237-Allele 1	411	0.00484	Locus-11	224	0.00014
D237-Allele 1	416	0.03872	Locus-11	228	0.19259
D237-Allele 1	420	0.05300	Locus-11	232	0.10754
D237-Allele 1	424	0.16505	Locus-11	236	0.00003
D237-Allele 1	429	0.01017	Locus-11	240	0.26316
D237-Allele 1	433	0.27541	Locus-11	244	0.00041
D237-Allele 1	436	0.00048	Locus-11	248	0.01024
D237-Allele 1	437	0.05711	Locus-11	252	0.00006
D237-Allele 1	441	0.05469	Locus-11	256	0.20367
D237-Allele 1	445	0.04017	Locus-11	260	0.00006
D237-Allele 1	449	0.06970	Locus-11	264	0.00087

APPENDIX (continued)

Locus	Allele	Frequency	Locus	Allele	Frequency
Locus-12	230	0.33116	Locus-14	294	0.13620
Locus-12	234	0.10994	Locus-14	298	0.00657
Locus-12	238	0.14960	Locus-14	302	0.03873
Locus-12	242	0.18259	Locus-14	306	0.00808
Locus-12	246	0.02344	Locus-15	260	0.18617
Locus-12	250	0.11928	Locus-15	262	0.33077
Locus-12	254	0.01664	Locus-15	264	0.00868
Locus-12	258	0.06734	Locus-15	266	0.01124
Locus-13	240	0.11642	Locus-15	268	0.02511
Locus-13	244	0.24408	Locus-15	270	0.18601
Locus-13	248	0.02605	Locus-15	272	0.25202
Locus-13	252	0.17822	Locus-16	270	0.00243
Locus-13	256	0.16224	Locus-16	272	0.23307
Locus-13	260	0.04907	Locus-16	274	0.17366
Locus-13	264	0.07767	Locus-16	276	0.00030
Locus-13	268	0.05972	Locus-16	278	0.13517
Locus-13	272	0.00278	Locus-16	280	0.00148
Locus-13	276	0.06325	Locus-16	282	0.00874
Locus-13	280	0.00009	Locus-16	284	0.30209
Locus-13	284	0.00061	Locus-16	286	0.14307
Locus-13	288	0.01981			
Locus-14	250	0.11427			
Locus-14	254	0.10100			
Locus-14	258	0.00054			
Locus-14	262	0.10672			
Locus-14	266	0.04113			
Locus-14	270	0.00071			
Locus-14	274	0.00173			
Locus-14	278	0.12635			
Locus-14	282	0.21178			
Locus-14	286	0.10605			
Locus-14	290	0.00015			

BIBLIOGRAPHY

Almudevar, A. & Field, C. 1999, "Estimation of single-generation sibling relationships based on DNA markers", *Journal of Agricultural Biological and Environmental Statistics*, vol. 4, no. 2, pp. 136-165.

Almudevar, A. 2007, "A graphical approach to relatedness inference", *Theoretical Population Biology*, vol. 71, no. 2, pp. 213-229.

Anderson, E. C. & Dunham, K. K. 2005, "SPIP 1.0: a program for simulating pedigrees and genetic data in age-structured populations", *Molecular Ecology Notes*, vol. 5, no. 2, pp. 459-461.

Anderson, E. C. & Thompson, E. A. 2002, "A model-based method for identifying species hybrids using multilocus genetic data", *Genetics*, vol. 160, no. 3, pp. 1217-1229.

Anderson, E. C. 2005, "An efficient Monte Carlo method for estimating N-e from temporally spaced samples using a coalescent-based likelihood", *Genetics*, vol. 170, no. 2, pp. 955-967.

Araki, H., Waples, R. S., Ardren, W. R., Cooper, B., & Blouin, M. S. 2007, "Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms", *Molecular Ecology*, vol. 16, no. 5, pp. 953-966.

Ardren, W. R. & Kapuscinski, A. R. 2003, "Demographic and genetic estimates of effective population size (N-e) reveals genetic compensation in steelhead trout", *Molecular Ecology*, vol. 12, no. 1, pp. 35-49.

Ashley, M. V., Caballero, I. C., Chaovalitwongse, W., Dasgupta, B., Govindan, P., Sheikh, S. I., & Berger-Wolf, T. Y. 2009, "KINALYZER, a computer program for reconstructing sibling groups", *Molecular Ecology Resources*, vol. 9, no. 4, pp. 1127-1131.

Balloux, F. 2001, "EASYPOP (version 1.7): A computer program for population genetics simulations", *Journal of Heredity*, vol. 92, no. 3, pp. 301-302.

Banks, M. A. & Eichert, W. 2000, "WHICHRUN (version 3.2): A computer program for population assignment of individuals based on multilocus genotype data", *Journal of Heredity*, vol. 91, no. 1, pp. 87-89.

Bateman, A. J. 1948, "Intra-Sexual Selection in Drosophila", *Heredity*, vol. 2, no. 3, pp. 349-368.

Beaumont, M. A. & Nichols, R. A. 1996, "Evaluating loci for use in the genetic analysis of population structure", *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 263, no. 1377, pp. 1619-1626.

Beaumont, M. A. 1999, "Detecting population expansion and decline using microsatellites", *Genetics*, vol. 153, no. 4, pp. 2013-2029.

Beaumont, M. A. 2003, "Estimation of population growth or decline in genetically monitored populations", *Genetics*, vol. 164, no. 3, pp. 1139-1160.

Beerli, P. 2006, "Comparison of Bayesian and maximum-likelihood inference of population genetic parameters", *Bioinformatics*, vol. 22, no. 3, pp. 341-345.

Belkhir, K., Borsa P., Chikhi, L., Raufaste N., & Bonhomme, F. 2004, "GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations", Laboratoire Genome, Populations, Interactions, CNRS UMR 5171, Universite de Montpellier II, Montpellier, France.

Belkhir, K., Castric, V., & Bonhomme, F. 2002, "IDENTIX, a software to test for relatedness in a population using permutation methods", *Molecular Ecology Notes*, vol. 2, no. 4, pp. 611-614.

Bell, G. 1980, "The Costs of Reproduction and Their Consequences", *American Naturalist*, vol. 116, no. 1, pp. 45-76.

Berg, O. K., Thronaes, E., & Bremset, G. 1998, "Energetics and survival of virgin and repeat spawning brown trout (*Salmo trutta*)", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 55, no. 1, pp. 47-53.

Berrigan, D. & Charnov, E. L. 1994, "Reaction Norms for Age and Size at Maturity in Response to Temperature - A Puzzle for Life Historians", *Oikos*, vol. 70, no. 3, pp. 474-478.

Berthier, P., Beaumont, M. A., Cornuet, J. M., & Luikart, G. 2002, "Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach", *Genetics*, vol. 160, no. 2, pp. 741-751.

Blanchfield, P. J. & Ridgway, M. S. 1999, "The cost of peripheral males in a brook trout mating system", *Animal Behaviour*, vol. 57, pp. 537-544.

Blanchfield, P. J. & Ridgway, M. S. 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*, vol. 62, no. 12, pp. 2694-2705.

Blanchfield, P. J., Ridgway, M. S., & Wilson, C. C. 2003, "Breeding success of male brook trout (*Salvelinus fontinalis*) in the wild", *Molecular Ecology*, vol. 12, no. 9, pp. 2417-2428.

Blouin, M. S. 2003, "DNA-based methods for pedigree reconstruction and kinship analysis in natural populations", *Trends in Ecology & Evolution*, vol. 18, no. 10, pp. 503-511.

Bohonak, A. J. 2002, "IBD (isolation by distance): A program for analyses of isolation by distance", *Journal of Heredity*, vol. 93, no. 2, pp. 153-154.

Boichard, D. "PEDIG: A Fortran package for pedigree analysis suited to large populations.", 7th World Congress on Genetics Applied to Livestock Production, Montpellier, 19-23 August, 2002, Paper 28-13.

Butler, K., Field, C., Herbinger, C. M., & Smith, B. R. 2004, "Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data", *Molecular Ecology*, vol. 13, no. 6, pp. 1589-1600.

Carvajal-Rodriguez, A. 2007, "FAMSPHERE: a computer program for parental allocation from known genotypic pools", *Molecular Ecology Notes*, vol. 7, no. 2, pp. 213-216.

Cercueil, A., Bellemain, E., & Manel, S. 2002, "PARENTE: Computer program for parentage analysis", *Journal of Heredity*, vol. 93, no. 6, pp. 458-459.

Charlesworth, B. 2009, "Effective population size and patterns of molecular evolution and variation", *Nature Reviews Genetics*, vol. 10, no. 3, pp. 195-205.

Charmantier, A. & Reale, D. 2005, "How do misassigned paternities affect the estimation of heritability in the wild?", *Molecular Ecology*, vol. 14, no. 9, pp. 2839-2850.

Choquet, R., Lebreton, J. D., Gimenez, O., Reboulet, A. M., & Pradel, R. 2009, "U-CARE: Utilities for performing goodness of fit tests and manipulating CApture-REcapture data", *Ecography*, vol. 32, no. 6, pp. 1071-1074.

Choquet, R., Reboulet, A. M., Pradel, R., Gimenez, O., & Lebreton, J. D. 2004, "M-SURGE: new software specifically designed for multistate capture-recapture models", *Animal Biodiversity and Conservation*, vol. 27, no. 1, pp. 207-221.

Clutton-Brock, T. 2007, "Sexual selection in males and females", *Science*, vol. 318, no. 5858, pp. 1882-1885.

Cockburn, A., Osmond, H. L., & Double, M. C. 2008, "Swingin' in the rain: condition dependence and sexual selection in a capricious world", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 605-612.

Cole, L. C. 1954, "The Population Consequences of Life History Phenomena", *Quarterly Review of Biology*, vol. 29, no. 2, pp. 103-137.

Coombs, J. A. 2010, *Reproduction in the Wild: The Effect of Individual Life History Strategies on Population Dynamics and Persistence* Ph.D Dissertation, University of Massachusetts, Amherst.

Coombs, J. A., Letcher, B. H., & Nislow, K. H. 2008, "CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs", *Molecular Ecology Resources*, vol. 8, no. 3, pp. 578-580.

Coombs, J. A., Letcher, B. H., & Nislow, K. H. 2010a, "PEDAGOG: Software for simulating eco-evolutionary population dynamics", *Molecular Ecology Resources*, vol. 10, no. 3, pp. 558-563.

Coombs, J. A., Letcher, B. H., & Nislow, K. H. 2010b, "PedAgree: Software to quantify error and assess accuracy and congruence for genetically reconstructed pedigree relationships", *Conservation Genetics Resources*, In Press.

Corander, J. & Marttinen, P. 2006, "Bayesian identification of admixture events using multilocus molecular markers", *Molecular Ecology*, vol. 15, no. 10, pp. 2833-2843.

Coulson, T., Benton, T. G., Lundberg, P., Dall, S. R. X., Kendall, B. E., & Gaillard, J. M. 2006, "Estimating individual contributions to population growth: evolutionary fitness in ecological time", *Proceedings of the Royal Society B-Biological Sciences*, vol. 273, no. 1586, pp. 547-555.

Crespi, B. J. 1989, "Causes of Assortative Mating in Arthropods", *Animal Behaviour*, vol. 38, pp. 980-1000.

Crnokrak, P. & Roff, D. A. 1999, "Inbreeding depression in the wild", *Heredity*, vol. 83, pp. 260-270.

Curry, R. A., Noakes, D. L. G., & Morgan, G. E. 1995, "Groundwater and the Incubation and Emergence of Brook Trout (*Salvelinus fontinalis*)", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 52, no. 8, pp. 1741-1749.

Curry, R. A., Powles, P. M., Gunn, J. M., & Liimatainen, V. A. 1991, "Emergence Chronology of Brook Charr, *Salvelinus fontinalis*, Alevins in An Acidic Stream", *Environmental Biology of Fishes*, vol. 31, no. 1, pp. 25-31.

Danzmann, R. G. 1997, "PROBMAX: A computer program for assigning unknown parentage in pedigree analysis from known genotypic pools of parents and progeny", *Journal of Heredity*, vol. 88, no. 4, p. 333.

Day, T. & Rowe, L. 2002, "Developmental thresholds and the evolution of reaction norms for age and size at life-history transitions", *American Naturalist*, vol. 159, no. 4, pp. 338-350.

Dickerson, B. R., Willson, M. F., Bentzen, P., & Quinn, T. P. 2004, "Size-assortative mating in salmonids: negative evidence for pink salmon in natural conditions", *Animal Behaviour*, vol. 68, pp. 381-385.

Dieringer, D. & Schlotterer, C. 2003, "Microsatellite Analyser (Msa): A Platform Independent Analysis Tool for Large Microsatellite Data Sets", *Molecular Ecology Notes*, vol. 3, no. 1, pp. 167-169.

Duchesne, P., Castric, T., & Bernatchez, L. 2005, "PASOS (parental allocation of singles in open systems): a computer program for individual parental allocation with missing parents", *Molecular Ecology Notes*, vol. 5, no. 3, pp. 701-704.

Duchesne, P., Godbout, M. H., & Bernatchez, L. 2002, "PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation", *Molecular Ecology Notes*, vol. 2, no. 2, pp. 191-193.

Einum, S. & Fleming, I. A. 1999, "Maternal effects of egg size in brown trout (*Salmo trutta*): norms of reaction to environmental quality", *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 266, no. 1433, pp. 2095-2100.

Einum, S. & Fleming, I. A. 2000a, "Highly fecund mothers sacrifice offspring survival to maximize fitness", *Nature*, vol. 405, no. 6786, pp. 565-567.

Einum, S. & Fleming, I. A. 2000b, "Selection against late emergence and small offspring in Atlantic salmon (*Salmo salar*)", *Evolution*, vol. 54, no. 2, pp. 628-639.

Einum, S. & Nislow, K. H. 2005, "Local-scale density-dependent survival of mobile organisms in continuous habitats: an experimental test using Atlantic salmon", *Oecologia*, vol. 143, no. 2, pp. 203-210.

Elliott, J. M. 1984, "Numerical Changes and Population Regulation in Young Migratory Trout (*Salmo trutta*) in A Lake District Stream, 1966-83", *Journal of Animal Ecology*, vol. 53, no. 1, pp. 327-350.

Emery, A. M., Wilson, I. J., Craig, S., Boyle, P. R., & Noble, L. R. 2001, "Assignment of paternity groups without access to parental genotypes: multiple mating and developmental plasticity in squid", *Molecular Ecology*, vol. 10, no. 5, pp. 1265-1278.

Esteve, M. 2005, "Observations of spawning behaviour in salmoninae: *Salmo, Oncorhynchus and Salvelinus*", *Reviews in Fish Biology and Fisheries*, vol. 15, no. 1-2, pp. 1-21.

Ewen, K. R., Bahlo, M., Treloar, S. A., Levinson, D. F., Mowry, B., Barlow, J. W., & Foote, S. J. 2000, "Identification and analysis of error types in high-throughput genotyping", *American Journal of Human Genetics*, vol. 67, no. 3, pp. 727-736.

Excoffier, L. & Heckel, G. 2006, "Computer programs for population genetics data analysis: a survival guide", *Nature Reviews Genetics*, vol. 7, no. 10, pp. 745-758.

Excoffier, L., Laval, G., & Schneider, S. 2005, "Arlequin ver. 3.0: An integrated software package for population genetics data analysis", *Evolutionary Bioinformatics Online*, vol. 1, pp. 47-50.

Falconer, D. S. & Mackay, T. F. C. 1996, *Introduction to quantitative genetics*, fourth edn, Longman, Harlow, England.

Falush, D., Stephens, M., & Pritchard, J. K. 2003, "Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies", *Genetics*, vol. 164, no. 4, pp. 1567-1587.

Felsenstein, J. 2004, "PHYLIP (Phylogeny Inference Package) version 3.6", Department of Genome Sciences, University of Washington, Seattle (Program available from: http://evolution.genetics.washington.edu/phylip.html).

Fleming, I. A. 1996, "Reproductive strategies of Atlantic salmon: Ecology and evolution", *Reviews in Fish Biology and Fisheries*, vol. 6, no. 4, pp. 379-416.

Foll, M. & Gaggiotti, O. E. 2005, "COLONISE: a computer program to study colonization processes in metapopulations", *Molecular Ecology Notes*, vol. 5, no. 3, pp. 705-707.

Foote, C. J. & Larkin, P. A. 1988, "The Role of Male Choice in the Assortative Mating of Anadromous and Non-Anadromous Sockeye Salmon (Oncorhynchus-Nerka)", *Behaviour*, vol. 106, pp. 43-62.

Foote, C. J. 1989, "Female Mate Preference in Pacific Salmon", *Animal Behaviour*, vol. 38, pp. 721-723.

Frankham, R. 1995, "Effective Population-Size Adult-Population Size Ratios in Wildlife - A Review", *Genetical Research*, vol. 66, no. 2, pp. 95-107.

Frankham, R. 2005, "Genetics and extinction", *Biological Conservation*, vol. 126, no. 2, pp. 131-140.

Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. 1999, "Do population size bottlenecks reduce evolutionary potential?", *Animal Conservation*, vol. 2, no. 4, pp. 255-260.

Franklin, I. R. & Frankham, R. 1998, "How large must populations be to retain evolutionary potential?", *Animal Conservation*, vol. 1, no. 1, pp. 69-70.

Fraser, D. J., Hansen, M. M., Ostergaard, S., Tessier, N., Legault, M., & Bernatchez, L. 2007, "Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems", *Molecular Ecology*, vol. 16, no. 18, pp. 3866-3889.

Gadgil, M. & Bossert, W. H. 1970, "Life Historical Consequences of Natural Selection", *The American Naturalist*, vol. 104, no. 935, pp. 1-24.

Gamito, S. 1998, "Growth models and their use in ecological modelling: an application to a fish population", *Ecological Modelling*, vol. 113, no. 1-3, pp. 83-94.

Garant, D., Fontaine, P. M., Good, S. P., Dodson, J. J., & Bernatchez, L. 2002, "The influence of male parental identity on growth and survival of offspring in Atlantic salmon (*Salmo salar*)", *Evolutionary Ecology Research*, vol. 4, no. 4, pp. 537-549.

Garant, D., Kruuk, L. E. B., McCleery, R. H., & Sheldon, B. C. 2004, "Evolution in a changing environment: A case study with great tit fledging mass", *American Naturalist*, vol. 164, no. 5, p. E115-E129.

Gerber, S., Chabrier, P., & Kremer, A. 2003, "FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers", *Molecular Ecology Notes*, vol. 3, no. 3, pp. 479-481.

Gerber, S., Mariette, S., Streiff, R., Bodenes, C., & Kremer, A. 2000, "Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis", *Molecular Ecology*, vol. 9, no. 8, pp. 1037-1048.

Glaubitz, J. C. 2004, "CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages", *Molecular Ecology Notes*, vol. 4, no. 2, pp. 309-310.

Goudet, J. 1995, "FSTAT (Version 1.2): A computer program to calculate F-statistics", *Journal of Heredity*, vol. 86, no. 6, pp. 485-486.

Grant, P. R. & Grant, B. R. 2008, "Pedigrees, assortative mating and speciation in Darwin's finches", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 661-668.

Gries, G. & Letcher, B. H. 2002, "Tag retention and survival of age-0 Atlantic salmon following surgical implantation with passive integrated transponder tags", *North American Journal of Fisheries Management*, vol. 22, no. 1, pp. 219-222.

Grossman, G. D., Ratajczak, R. E., Wagner, C. M., & Petty, J. T. 2010, "Dynamics and regulation of the southern brook trout (*Salvelinus fontinalis*) population in an Appalachian stream", *Freshwater Biology*, vol. 55, no. 7, pp. 1494-1508.

Grover, M. C. 2005, "Changes in size and age at maturity in a population of kokanee Oncorhynchus nerka during a period of declining growth conditions", *Journal of Fish Biology*, vol. 66, no. 1, pp. 122-134.

Haig, D. & Bergstrom, C. T. 1995, "Multiple Mating, Sperm Competition and Meiotic Drive", *Journal of Evolutionary Biology*, vol. 8, no. 3, pp. 265-282.

Hairston, N. G., Ellner, S. P., Geber, M. A., Yoshida, T., & Fox, J. A. 2005, "Rapid evolution and the convergence of ecological and evolutionary time", *Ecology Letters*, vol. 8, no. 10, pp. 1114-1127.

Hanski, I. & Saccheri, I. 2006, "Molecular-level variation affects population growth in a butterfly metapopulation", *Plos Biology*, vol. 4, no. 5, pp. 719-726.

Hardy, O. J. & Vekemans, X. 2002, "SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels", *Molecular Ecology Notes*, vol. 2, no. 4, pp. 618-620.

Harvey, P. H. & Zammuto, R. M. 1985, "Patterns of Mortality and Age at 1St Reproduction in Natural-Populations of Mammals", *Nature*, vol. 315, no. 6017, pp. 319-320.

Hedgecock, D. & Eichert, W. 1999, "WHICHPARENTS (version 1.0): A windows application for determining the most likely parents of offspring using multilocus genotype data", Computer software distributed from: http://www.bml.ucdavis.edu/whichparents.html.

Hedrick, P. 2005, "Large variance in reproductive success and the N-e/N ratio", *Evolution*, vol. 59, no. 7, pp. 1596-1599.

Hendry, A. P. & Stearns, S. C. 2004, *Evolution Illuminated: Salmon and their relatives* Oxford University Press, Oxford.

Hey, J. & Nielsen, R. 2004, "Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D-persimilis", *Genetics*, vol. 167, no. 2, pp. 747-760.

Hey, J. & Nielsen, R. 2007, "Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 8, pp. 2785-2790.

Holm, S. 1979, "A Simple Sequentially Rejective Multiple Test Procedure", *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65-70.

Hudy, M., Coombs, J. A., Nislow, K. H., & Letcher, B. H. 2010, "Dispersal and within-stream spatial population structure of brook trout revealed by pedigree reconstruction analysis", *Transactions of the American Fisheries Society*, In Press.

Hudy, M., Downey, D. M., & Bowman, D. W. 2000, "Successful Restoration of an Acidified Native Brook Trout Stream through Mitigation with Limestone Sand", *North American Journal of Fisheries Management*, vol. 20, no. 2, pp. 453-466.

Hudy, M., Thieling, T. M., Gillespie, N., & Smith, E. P. 2008, "Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the Eastern United States", *North American Journal of Fisheries Management*, vol. 28, no. 4, pp. 1069-1085.

Hutchings, J. A. 1994, "Age-Specific and Size-Specific Costs of Reproduction Within Populations of Brook Trout, *Salvelinus fontinalis*", *Oikos*, vol. 70, no. 1, pp. 12-20.

Hutchings, J. A. 1996, "Adaptive phenotypic plasticity in brook trout, *Salvelinus fontinalis*, life histories", *Ecoscience*, vol. 3, no. 1, pp. 25-32.

Hutchings, J. A. 2006, "Survival consequences of sex-biased growth and the absence of a growth-mortality trade-off", *Functional Ecology*, vol. 20, no. 2, pp. 347-353.

Hutchings, J. A., Pickle, A., McGregor-Shaw, C. R., & Poirier, L. 1999, "Influence of sex, body size, and reproduction on overwinter lipid depletion in brook trout", *Journal of Fish Biology*, vol. 55, no. 5, pp. 1020-1028.

Jacob, A., Nussle, S., Britschgi, A., Evanno, G., Muller, R., & Wedekind, C. 2007, "Male dominance linked to size and age, but not to 'good genes' in brown trout (*Salmo trutta*)", *BMC Evolutionary Biology*, vol. 7.

Jenkins, T. M., Diehl, S., Kratz, K. W., & Cooper, S. D. 1999, "Effects of population density on individual growth of brown trout in streams", *Ecology*, vol. 80, no. 3, pp. 941-956.

Jones, A. G. & Ardren, W. R. 2003, "Methods of parentage analysis in natural populations", *Molecular Ecology*, vol. 12, no. 10, pp. 2511-2523.
Jones, A. G. 2001, "GERUD1.0: a computer program for the reconstruction of parental genotypes from progeny arrays using multilocus DNA data", *Molecular Ecology Notes*, vol. 1, no. 3, pp. 215-218.

Jones, A. G. 2005, "GERUD 2.0: a computer program for the reconstruction of parental genotypes from half-sib progeny arrays with known or unknown parents", *Molecular Ecology Notes*, vol. 5, no. 3, pp. 708-711.

Jones, B., Grossman, G. D., Walsh, D. C. I., Porter, B. A., Avise, J. C., & Fiumera, A. C. 2007, "Estimating differential reproductive success from nests of related individuals, with application to a study of the mottled sculpin, Cottus bairdi", *Genetics*, vol. 176, no. 4, pp. 2427-2439.

Jones, M. W. & Hutchings, J. A. 2001, "The influence of male parr body size and mate competition on fertilization success and effective population size in Atlantic salmon", *Heredity*, vol. 86, pp. 675-684.

Jones, M. W. & Hutchings, J. A. 2002, "Individual variation in Atlantic salmon fertilization success: Implications for effective population size", *Ecological Applications*, vol. 12, no. 1, pp. 184-193.

Jorde, P. E. & Ryman, N. 1995, "Temporal Allele Frequency Change and Estimation of Effective Size in Populations with Overlapping Generations", *Genetics*, vol. 139, no. 2, pp. 1077-1090.

Jorde, P. E. & Ryman, N. 2007, "Unbiased estimator for genetic drift and effective population size", *Genetics*, vol. 177, no. 2, pp. 927-935.

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. 2007, "Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment", *Molecular Ecology*, vol. 16, no. 5, pp. 1099-1106.

Kalinowski, S. T., Wagner, A. P., & Taper, M. L. 2006, "ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship", *Molecular Ecology Notes*, vol. 6, no. 2, pp. 576-579.

Keller, L. F. & Waller, D. M. 2002, "Inbreeding effects in wild populations", *Trends in Ecology & Evolution*, vol. 17, no. 5, pp. 230-241.

Keller, L. F., Reid, J. M., & Arcese, P. 2008, "Testing evolutionary models of senescence in a natural population: age and inbreeding effects on fitness components in song sparrows", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 597-604.

Kimura, M. & Crow, J. F. 1964, "Number of Alleles That Can be Maintained in Finite Population", *Genetics*, vol. 49, no. 4, p. 725-&.

King, T. L., Julian, S. E., Coleman, R. L., & Burnham-Curtis, M. K. 2003, "Isolation and characterization of novel tri- and tetra-nucleotide microsatellite DNA markers for brook trout *Salvelinus fontinalis*", Genbank submission numbers: AY168187, AY168192, AY 168193, AY168194, AY168195, AY168197, AY168199. Available at http://www.ncbi.nlm.nih.gov/nucleotide/.

King, T. L., Eackles, M. S., & Letcher, B. H. 2005, "Microsatellite DNA markers for the study of Atlantic salmon (*Salmo salar*) kinship, population structure, and mixed-fishery analyses", *Molecular Ecology Notes*, vol. 5, no. 1, pp. 130-132.

Kinnison, M. T. & Hairston, N. G. 2007, "Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence", *Functional Ecology*, vol. 21, no. 3, pp. 444-454.

Konovalov, D. A., Manning, C., & Henshaw, M. T. 2004, "KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers", *Molecular Ecology Notes*, vol. 4, no. 4, pp. 779-782.

Korfiatis, K., Papatheodorou, E., Stamou, G. P., & Paraskevopoulous, S. 1999, "An investigation of the effectiveness of computer simulation programs as tutorial tools for teaching population ecology at university", *International Journal of Science Education*, vol. 21, no. 12, pp. 1269-1280.

Kruuk, L. E. B. & Hill, W. G. 2008, "Introduction. Evolutionary dynamics of wild populations: the use of long-term pedigree data", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 593-596.

Kruuk, L. E. B. 2004, "Estimating genetic parameters in natural populations using the 'animal model'", *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, vol. 359, no. 1446, pp. 873-890.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., Brotherstone, S., Guinness, F., & Clutton-Brock, T. 2002, "Antler size in red deer: Heritability and selection but no evolution", *Evolution*, vol. 56, no. 8, pp. 1683-1695.

Kuhner, M. K. 2006, "LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters", *Bioinformatics*, vol. 22, no. 6, pp. 768-770.

Labonne, J., Augery, M., Parade, M., Brinkert, S., Prevost, E., Heland, M., & Beall, E. 2009, "Female preference for male body size in brown trout, *Salmo trutta*: is big still fashionable?", *Animal Behaviour*, vol. 77, no. 1, pp. 129-137.

Letcher, B. H., Gries, G., & Juanes, F. 2002, "Survival of stream-dwelling Atlantic salmon: Effects of life history variation, season, and age", *Transactions of the American Fisheries Society*, vol. 131, no. 5, pp. 838-854.

Letcher, B. H., Nislow, K. H., Coombs, J. A., O'Donnell, M. J., & Dubreuil, T. L. 2007, "Population Response to Habitat Fragmentation in a Stream-Dwelling Brook Trout Population", *PLoS ONE*, vol. 2, no. 11, p. e1139.

Lewis, P. O. & Zaykin, D. 2001, "Genetic Data Analysis: Computer program for the analysis of allelic data (Version 1.1)", Free program distributed by the authors over the internet from: http://lewis.eeb.uconn.edu/lewishome/software.html.

Lindstrom, J. & Kokko, H. 2002, "Cohort effects and population dynamics", *Ecology Letters*, vol. 5, no. 3, pp. 338-344.

Lobon-Cervia, J. & Rincon, P. A. 2004, "Environmental determinants of recruitment and their influence on the population dynamics of stream-living brown trout (*Salmo trutta*)", *Oikos*, vol. 105, no. 3, pp. 641-646.

Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. 2010, "Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches", *Conservation Genetics*, vol. 11, no. 2, pp. 355-373.

Manoukis, N. C. 2007, "FORMATOMATIC: A program for converting diploid allelic data between common formats for population genetic analysis", *Molecular Ecology Notes*, vol. 7, no. 4, pp. 592-593.

Marshall, E. & Randhir, T. 2008, "Effect of climate change on watershed system: a regional analysis", *Climatic Change*, vol. 89, no. 3-4, pp. 263-280.

Marshall, T. C., Slate, J., Kruuk, L. E. B., & Pemberton, J. M. 1998, "Statistical confidence for likelihood-based paternity inference in natural populations", *Molecular Ecology*, vol. 7, no. 5, pp. 639-655.

Meyer, K. 2006, "WOMBAT - A program for mixed model analyses by restricted maximum likelihood", User notes. Animal Genetics and Breeding Unit, Armidale, 58pp.

Miller, M. P. 1997, "Tools For Population Genetic Analyses (TFPGA): A windows program for the analysis of allozyme and molecular population genetic data", Computer software distributed from: http://www.marksgeneticsoftware.net.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., & Cavalli-Sforza, L. L. 1996, "MICROSAT (version 1.5b): A computer program for calculating various statistics on microsatellite allele data", Computer software distributed from: http://hpgl.stanford.edu/projects/microsat/.

Morita, K., Morita, S. H., Fukuwaka, M., & Matsuda, H. 2005, "Rule of age and size at maturity of chum salmon (Oncorhynchus keta): implications of recent trends among Oncorhynchus spp", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 62, no. 12, pp. 2752-2759.

Morita, K. & Takashima, Y. 1998, “Effect of female size on fecundity and egg size in white-spotted charr: comparison between sea-run and resident forms”, *Journal of Fish Biology*, vol. 53, no. 5, pp. 1140-1142.

Morrissey, M. B., Wilson, A. J., Pemberton, J. M., & Ferguson, M. M. 2007, "A framework for power and sensitivity analyses for quantitative genetic studies of natural populations, and case studies in Soay sheep (Ovis aries)", *Journal of Evolutionary Biology*, vol. 20, no. 6, pp. 2309-2321.

Nagylaki, T. 1979, "Island Model with Stochastic Migration", *Genetics*, vol. 91, no. 1, pp. 163-176.

Neumaier, A. & Groeneveld, E. 1998, "Restricted maximum likelihood estimation of covariances in sparse linear models", *Genetics Selection Evolution*, vol. 30, no. 1, pp. 3-26.

Newman, D. & Pilson, D. 1997, "Increased probability of extinction due to decreased genetic effective population size: Experimental populations of Clarkia pulchella", *Evolution*, vol. 51, no. 2, pp. 354-362.

Nielsen, J. P. 1999, Record extension and streamflow statistics for the Pleasant River, Maine, USGS, http://me.water.usgs.gov/reports/finalreport.pdf, SuDOoc I 19.42/4:99-4078.

Nislow, K. H. & Lowe, W. H. 2003, "Influences of logging history and stream pH on brook trout abundance in first-order streams in New Hampshire", *Transactions of the American Fisheries Society*, vol. 132, no. 1, pp. 166-171.

Nislow, K. H., Sepulveda, A. J., & Folt, C. L. 2004, "Mechanistic linkage of hydrologic regime to summer growth of age-0 Atlantic salmon", *Transactions of the American Fisheries Society*, vol. 133, no. 1, pp. 79-88.

O'Connor, K. D., Marr, A. B., Arcese, P., Keller, L. F., Jeffery, K. J., & Bruford, M. W. 2006, "Extra-pair fertilization and effective population size in the song sparrow Melospiza melodia", *Journal of Avian Biology*, vol. 37, no. 6, pp. 572-578.

Ohta, T. & Kimura, M. 1973, "Model of Mutation Appropriate to Estimate Number of Electrophoretically Detectable Alleles in A Finite Population", *Genetical Research*, vol. 22, no. 2, pp. 201-204.

O'Reilly, P. T., Herbinger, C., & Wright, J. M. 1998, "Analysis of parentage determination in Atlantic salmon (*Salmo salar*) using microsatellites", *Animal Genetics*, vol. 29, no. 5, pp. 363-370.

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*, vol. 17, no. 15, pp. 3428-3447.

Peel, D., Ovenden, J. R., & Peel, S. L. NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries. 2004.

Pelletier, F., Clutton-Brock, T., Pemberton, J., Tuljapurkar, S., & Coulson, T. 2007, "The evolutionary demography of ecological change: Linking trait variation and population growth", *Science*, vol. 315, no. 5818, pp. 1571-1574.

Pemberton, J. M. 2008, "Wild pedigrees: the way forward", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 613-621.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. 2004, "GENECLASS2: A software for genetic assignment and first-generation migrant detection", *Journal of Heredity*, vol. 95, no. 6, pp. 536-539.

Piry, S., Luikart, G., & Cornuet, J. M. 1999, "BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data", *Journal of Heredity*, vol. 90, no. 4, pp. 502-503.

Price, T. A. R., Hurst, G. D. D., & Wedell, N. 2010, "Polyandry Prevents Extinction", *Current Biology*, vol. 20, no. 5, pp. 471-475.

Rannala, B. & Mountain, J. L. 1997, "Detecting immigration by using multilocus genotypes", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 17, pp. 9197-9201.

Raymond, M. & Rousset, F. 1995, "Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism", *Journal of Heredity*, vol. 86, no. 3, pp. 248-249.

Reed, D. H. 2005, "Relationship between population size and fitness", *Conservation Biology*, vol. 19, no. 2, pp. 563-568.

Rice, W. R. 1989, "Analyzing Tables of Statistical Tests", *Evolution*, vol. 43, no. 1, pp. 223-225.

Rose, K. A., Cowan, J. H., Winemiller, K. O., Myers, R. A., & Hilborn, R. 2001, "Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis", *Fish and Fisheries*, vol. 2, no. 4, pp. 293-327.

Rose, M. R. & Charlesworth, B. 1981, "Genetics of Life-History in Drosophila-Melanogaster .2. Exploratory Selection Experiments", *Genetics*, vol. 97, no. 1, pp. 187-196.

Rousset, F. 2008, "GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux", *Molecular Ecology Resources*, vol. 8, no. 1, pp. 103-106.

Rowe, D. K., Thorpe, J. E., & Shanks, A. M. 1991, "Role of Fat Stores in the Maturation of Male Atlantic Salmon (*Salmo salar*) Parr", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 48, no. 3, pp. 405-413.

Saccheri, I. & Hanski, I. 2006, "Natural selection and population dynamics", *Trends in Ecology & Evolution*, vol. 21, no. 6, pp. 341-347.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. 1998, "Inbreeding and extinction in a butterfly metapopulation", *Nature*, vol. 392, no. 6675, pp. 491-494.

Saenz-Agudelo, P., Jones, G. P., Thorrold, S. R., & Planes, S. 2009, "Estimating connectivity in marine populations: an empirical evaluation of assignment tests and parentage analysis under different gene flow scenarios", *Molecular Ecology*, vol. 18, no. 8, pp. 1765-1776.

Sieberts, S. K., Wijsman, E. M., & Thompson, E. A. 2002, "Relationship inference from trios of individuals, in the presence of typing error", *American Journal of Human Genetics*, vol. 70, no. 1, pp. 170-180.

Smith, B. R., Herbinger, C. M., & Merry, H. R. 2001, "Accurate partition of individuals into full-sib families from genetic data without parental information", *Genetics*, vol. 158, no. 3, pp. 1329-1338.

Snucins, E. J., Curry, R. A., & Gunn, J. M. 1992, "Brook Trout (*Salvelinus fontinalis*) Embryo Habitat and Timing of Alevin Emergence in A Lake and A Stream", *Canadian Journal of Zoology-Revue Canadienne de Zoologie*, vol. 70, no. 3, pp. 423-427.

Stearns, S. C. & Koella, J. C. 1986, "The Evolution of Phenotypic Plasticity in Life-History Traits - Predictions of Reaction Norms for Age and Size at Maturity", *Evolution*, vol. 40, no. 5, pp. 893-913.

Stearns, S. C. 1992, *The evolution of life histories* Oxford University Press, Oxford.

Svedin, N., Wiley, C., Veen, T., Gustafsson, L., & Qvarnstrom, A. 2008, "Natural and sexual selection against hybrid flycatchers", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 735-744.

Szulkin, M. & Sheldon, B. C. 2008, "Dispersal as a means of inbreeding avoidance in a wild bird population", *Proceedings of the Royal Society B-Biological Sciences*, vol. 275, no. 1635, pp. 703-711.

Taggart, J. B., McLaren, I. S., Hay, D. W., Webb, J. H., & Youngson, A. F. 2001, "Spawning success in Atlantic salmon (*Salmo salar* L.): a long-term DNA profiling-based study conducted in a natural stream", *Molecular Ecology*, vol. 10, no. 4, pp. 1047-1060.

Theriault, V., Garant, D., Bernatchez, L., & Dodson, J. J. 2007, "Heritability of life-history tactics and genetic correlation with body size in a natural population of brook charr (*Salvelinus fontinalis*)", *Journal of Evolutionary Biology*, vol. 20, no. 6, pp. 2266-2277.

Thorpe, J. E. 1994, "Reproductive strategies in Atlantic salmon, *Salmo salar* L", *Aquaculture & Fisheries Management*, vol. 25, no. 1, pp. 77-87.

Valiere, N. 2002, "GIMLET: a computer program for analysing genetic individual identification data", *Molecular Ecology Notes*, vol. 2, no. 3, pp. 377-379.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. 2004, "MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data", *Molecular Ecology Notes*, vol. 4, no. 3, pp. 535-538.

Vincenzi, S., Crivelli, A. J., Jesensek, D., & De Leo, G. A. 2008, "The role of density-dependent individual growth in the persistence of freshwater salmonid populations", *Oecologia*, vol. 156, no. 3, pp. 523-534.

Vincenzi, S., Crivelli, A. J., Jesensek, D., Rubin, J. F., & De Leo, G. A. 2007, "Density-dependent individual growth of marble trout (*Salmo marmoratus*) in the Soca and Idrijca river basins, Slovenia", *Hydrobiologia*, vol. 583, pp. 57-68.

Vladykov, V. D. 1956, "Fecundity of Wild Speckled Trout (*Salvelinus fontinalis*) in Quebec Lakes", *Journal of the Fisheries Research Board of Canada*, vol. 13, no. 6, p. 799-841.

Wade, M. J. & Arnold, S. J. 1980, "The Intensity of Sexual Selection in Relation to Male Sexual-Behavior, Female Choice, and Sperm Precedence", *Animal Behaviour*, vol. 28, no. MAY, pp. 446-461.

Wang, J. & Santure, A. W. 2009, "Parentage and Sibship Inference From Multilocus Genotype Data Under Polygamy", *Genetics*, vol. 181, no. 4, pp. 1579-1594.

Wang, J. 2007, "Parentage and sibship exclusions: higher statistical power with more family members", *Heredity*, vol. 99, no. 2, pp. 205-217.

Wang, J. L. & Whitlock, M. C. 2003, "Estimating effective population size and migration rates from genetic samples over space and time", *Genetics*, vol. 163, no. 1, pp. 429-446.

Wang, J. L. 2001, "A pseudo-likelihood method for estimating effective population size from temporally spaced samples", *Genetical Research*, vol. 78, no. 3, pp. 243-257.

Wang, J. L. 2004, "Sibship reconstruction from genetic data with typing errors", *Genetics*, vol. 166, no. 4, pp. 1963-1979.

Waples, R. S. & Do, C. 2008, "LDNE: a program for estimating effective population size from data on linkage disequilibrium", *Molecular Ecology Resources*, vol. 8, no. 4, pp. 753-756.

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", *Evolutionary Applications*, vol. 3, no. 3, pp. 244-262.

Waples, R. S. 2005, "Genetic estimates of contemporary effective population size: to what time periods do the estimates apply?", *Molecular Ecology*, vol. 14, no. 11, pp. 3335-3352.

Waples, R. S. 2006, "A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci", *Conservation Genetics*, vol. 7, no. 2, pp. 167-184.

Waples, R. S., Masuda, M., & Pella, J. 2007, "SALMONNb: a program for computing cohort-specific effective population sizes (N-b) in Pacific salmon and other semelparous species using the temporal method", *Molecular Ecology Notes*, vol. 7, no. 1, pp. 21-24.

Wattier, R., Engel, C. R., Saumitou-Laprade, P., & Valero, M. 1998, "Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in Gracilaria gracilis (Rhodophyta)", *Molecular Ecology*, vol. 7, no. 11, pp. 1569-1573.

White, G. C. & Burnham, K. P. 1999, "Program MARK: survival estimation from populations of marked animals", *Bird Study*, vol. 46, pp. 120-139.

Wilbur, H. M. & Collins, J. P. 1973, "Ecological Aspects of Amphibian Metamorphosis", *Science*, vol. 182, no. 4119, pp. 1305-1314.

Wilmer, J. W., Allen, P. J., Pomeroy, P. P., Twiss, S. D., & Amos, W. 1999, "Where have all the fathers gone? An extensive microsatellite analysis of paternity in the grey seal (Halichoerus grypus)", *Molecular Ecology*, vol. 8, no. 9, pp. 1417-1429.

Wilson, A. J. & Ferguson, M. M. 2002, "Molecular pedigree analysis in natural populations of fishes: approaches, applications, and practical considerations", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 59, no. 10, pp. 1696-1707.

Wilson, A. J., Pemberton, J. M., Pilkington, J. G., Clutton-Brock, T. H., & Kruuk, L. E. B. 2009, "Trading offspring size for number in a variable environment: selection on reproductive investment in female Soay sheep", *Journal of Animal Ecology*, vol. 78, no. 2, pp. 354-364.

Wilson, G. A. & Rannala, B. 2003, "Bayesian inference of recent migration rates using multilocus genotypes", *Genetics*, vol. 163, no. 3, pp. 1177-1191.
Wilson, I. J., Weale, M. E., & Balding, D. J. 2003, "Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities", *Journal of the Royal Statistical Society Series A-Statistics in Society*, vol. 166, pp. 155-188.

Winemiller, K. O. & Rose, K. A. 1992, "Patterns of Life-History Diversification in North-American Fishes - Implications for Population Regulation", *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 49, no. 10, pp. 2196-2218.

Wofford, J. E. B., Gresswell, R. E., & Banks, M. A. 2005, "Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout", *Ecological Applications*, vol. 15, no. 2, pp. 628-637.

Wright, S. 1931, "Evolution in Mendelian populations", *Genetics*, vol. 16, no. 2, pp. 0097-0159.

Xu, C., Letcher, B. H., & Nislow, K. H. 2010, “Context-specific influence of water temperature on brook trout growth rates in the field”, *Freshwater Biology*, In Press.