

A Model for Estimating the Minimum Number of Offspring to Sample in Studies of Reproductive Success

JOSEPH H. ANDERSON, ERIC J. WARD, AND STEPHANIE M. CARLSON

From the School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195. Joseph H. Anderson and Eric J. Ward are now at NOAA Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, WA 98112. Stephanie M. Carlson is now at Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA.

Address correspondence to Joseph H. Anderson at the address above, or e-mail: joe.anderson@noaa.gov.

Abstract

Molecular parentage permits studies of selection and evolution in fecund species with cryptic mating systems, such as fish, amphibians, and insects. However, there exists no method for estimating the number of offspring that must be assigned parentage to achieve robust estimates of reproductive success when only a fraction of offspring can be sampled. We constructed a 2-stage model that first estimated the mean (μ) and variance (ν) in reproductive success from published studies on salmonid fishes and then sampled offspring from reproductive success distributions simulated from the μ and ν estimates. Results provided strong support for modeling salmonid reproductive success via the negative binomial distribution and suggested that few offspring samples are needed to reject the null hypothesis of uniform offspring production. However, the sampled reproductive success distributions deviated significantly (χ^2 goodness-of-fit test p value < 0.05) from the known simulated reproductive success distribution at rates often > 0.05 and as high as 0.24, even when hundreds of offspring were assigned parentage. In general, reproductive success patterns were less accurate when offspring were sampled from cohorts with larger numbers of parents and greater variance in reproductive success. Our model can be reparameterized with data from other species and will aid researchers in planning reproductive success studies by providing explicit sampling targets required to accurately assess reproductive success.

Key words: natural selection, negative binomial, parentage, pedigree, power, salmonid, sexual selection

Studies of reproductive success directly measure the demographic contribution of each individual in a population to the next generation, and this approach has revealed much about the processes of selection and evolution. In taxa, such as birds and mammals, it is possible to estimate reproductive success via direct observation of mating, birthing, or parental care. Traditionally, these organisms have dominated studies of individual reproductive success (e.g., Clutton-Brock 1988) because acquiring high-quality data across multiple years is possible via direct observation. However, molecular parentage techniques have extended studies of reproductive success to natural populations of organisms with more cryptic mating systems, such as fish and amphibians (e.g., Jones and Ardren 2003). A key theme emerging from work on reproductive success of these species is that a significant proportion of breeders, particularly males, produce no offspring (Jones et al. 2002; Seamons et al. 2007; Tatarenkov et al. 2008).

A thorough analysis of reproductive success requires a complete census of breeding individuals in a population. This is often difficult to achieve, but is possible if breeders can be intercepted during migration to breeding areas (Seamons et al. 2007) or if they are located in small, isolated habitats (Jones et al. 2002; Tatarenkov et al. 2008). Incomplete sampling of parents or sampling only clutches of offspring has led to important insights into the behavior and reproductive ecology of fish (reviewed by Avise et al. 2002) and amphibians (e.g., Gopurenko et al. 2006; Liebgold et al. 2006) but unfortunately cannot provide a rigorous evaluation of reproductive success without considering all parents, including those that produce no offspring. Thus, complete sampling of parents extends the analysis beyond genetic mating patterns and permits estimation of natural selection on morphological and life-history traits, such as body size, breeding date, and alternative life-history strategies (Jones et al. 2002; Garant et al. 2003; Seamons et al. 2007).

In addition to extensive collection of parental tissues, researchers studying reproductive success must determine how to quantify fitness—specifically, which life-history stage of offspring should be sampled. Adult offspring will be less numerous than juvenile offspring due to accumulated lifetime mortality and thus might be censused completely. However, collecting adult offspring might be logistically difficult for species that disperse long distances from the natal site or for those that mature slowly or at different ages. For these species, sampling of embryonic or juvenile offspring may be desirable because they are accessible immediately after breeding. Complete sampling of juveniles may be impossible, however. Many species of fish, amphibians, and insects produce large clutches of offspring that are prohibitively numerous (i.e., expensive) to genotype in entirety. Juvenile offspring may also be elusive (e.g., fish in a stream) such that significant effort and resources must be devoted to acquiring samples. In these cases, only a portion of the offspring population can be realistically sampled.

Researchers must therefore consider the number of juvenile offspring samples that must be collected and assigned to individual parents to adequately represent the true distribution of reproductive success in the wild. Unfortunately, current statistical approaches in parentage analyses have not addressed this crucial question. Previous work has addressed sample sizes required to differentiate reproductive success between 2 groups but not among individuals (Hinrichsen 2003). Others have addressed the minimum samples taken from a single nest or clutch to describe the genetic mating system (DeWoody et al. 2000; Neff et al. 2000), but these approaches do not require exhaustive sampling of parental genotypes and therefore fail to consider the parents that produce no offspring. There is therefore a need to develop statistical tools to help researchers determine minimum offspring sample sizes for analysis of reproductive success in taxa that produce large numbers of offspring. Indeed, 2 recent studies addressing reproductive success in amphibians and fish genotyped virtually the entire offspring population ($n = 862$, Gopurenko et al. 2007; $n = 1493$, Tataronov et al. 2008) rather than reduce sample size and discuss sampling confidence.

In this study, we employ a Bayesian statistical approach to estimate the number of offspring that must be assigned to parents to adequately represent the true pattern of reproductive success. Our model is parameterized based on data from published research evaluating reproductive success within a taxonomic family with similar mating behavior, the salmonid fishes (Esteve 2005). These data display common features of reproductive success studies across all species of amphibians and fish: The distribution of reproductive success is highly skewed, with a large fraction of adults producing no offspring. The model evaluates the effects of varying the mean (μ) and variance (ν) of reproductive success, as well as the number of parents per sex (n_{par}), on type I and type II error rates. Our objective is to aid researchers in planning field

sampling of juvenile progeny so that parentage analysis will provide statistically robust examinations of reproductive success. The model is also useful post hoc to estimate sampling success given empirically observed data. We do not consider attributes of the genetic data (number of loci, polymorphism of loci) or parental sampling in the statistical confidence of individual assignments, as these factors have been evaluated elsewhere (Marshall et al. 1998; Bernatchez and Duchesne 2000).

Methods

We chose to work with fishes of the family salmonidae because these species have been a focus of previous research on reproductive success on naturally breeding populations. Salmonid fishes are amenable to studies of reproductive success because a complete census can often be taken of all breeding adults during migrations to spawning grounds. Published studies included in our analysis fulfilled 4 criteria: 1) They were based on naturally spawning populations (excluding studies of laboratory or experimental stream channels), 2) reproductive success was evaluated via stream sampling of juvenile fish, 3) offspring were assigned to individual parents (and not categories, such as wild vs. hatchery), and 4) studies included breeders that failed to reproduce in reported means and variances of reproductive success. We found 4 studies with a total of 18 cohorts (9 per sex, Table 1) that fulfilled these criteria for populations of Atlantic salmon *Salmo salar* (Garant et al. 2001), steelhead trout *Oncorhynchus mykiss* (Seamons et al. 2004), coho salmon *O. kisutch* (Ford et al. 2006), and Chinook salmon *O. tshawytscha* (Baumsteiger et al. 2008). Ford et al. (2006) reported separate reproductive success values for hatchery and wild-origin parents, but we pooled these data using a pooled variance estimate (Zar 1999, p. 124). Reproductive success values published by Seamons et al. (2004) included only breeders that produced at least one offspring, so we obtained the raw data directly from the authors in order to include spawners that produced no offspring.

Our model was composed of 2 stages. In the first stage, we used published values of reproductive success from samples (Table 1) to estimate the distributions of true unknown population means and variances in our study taxa. In the second stage, we used the estimated parameter values from stage one to parameterize an offspring assignment model that recorded the number of offspring that must be assigned parents in order to accurately represent the distribution of reproductive success. Our model is thus generic with respect to sex: Data from both sexes were used to estimate a single dataset of true unknown means and variances.

Stage One: Parameter Estimation

For each of the 9 cohorts in our analysis, we allowed males and females to have different reproductive success parameters. Each dataset includes n_{par} parents of a given

Table 1 Summary of studies used to parameterize the reproductive success model

Species	Sex	Brood Year	ID	Model inputs						Estimated type I error rate	Reference
				n_{par}	n	$\hat{\mu}$	$\hat{\nu}$	f_{min}	f_{max}		
Atlantic salmon	Male	1995	a	41	593	14.5	124.8	0.00539	0.146	0.0200	Garant et al. (2001)
	Female		b	35	593	16.9	151.0	0.00539	0.146	0.0189	Garant et al. (2001)
Steelhead trout	Male	1997	c	11	80	7.27	49.8	0.00214	0.0580	0.0466	Seamons et al. (2004)
	Female		d	20	118	5.90	34.7	0.00315	0.0856	0.0431	Seamons et al. (2004)
	Male	1998	e	35	70	2.00	25.4	0.00208	0.0564	0.0962	Seamons et al. (2004)
	Female		f	18	111	6.17	46.0	0.00330	0.0895	0.0461	Seamons et al. (2004)
	Male	1999	g	25	90	3.60	48.3	0.00178	0.0484	0.0917	Seamons et al. (2004)
	Female		h	27	146	5.41	47.0	0.00289	0.0785	0.0547	Seamons et al. (2004)
	Male	2000	i	64	136	2.13	14.6	0.000969	0.0263	0.132	Seamons et al. (2004)
	Female		j	75	238	3.17	16.4	0.00170	0.0460	0.0771	Seamons et al. (2004)
Coho salmon	Male	2000	k	418	216	0.51	2.05	0.00243	0.0147	0.239	Ford et al. (2006)
	Female		l	389	216	0.55	2.20	0.00243	0.0147	0.237	Ford et al. (2006)
	Male	2001	m	522	345	0.66	3.44	0.00293	0.0178	0.234	Ford et al. (2006)
	Female		n	435	383	0.88	6.21	0.00326	0.0197	—	Ford et al. (2006)
Chinook salmon	Male	2002	o	63	53	0.84	3.46	0.00739	0.00820	0.154	Baumsteiger et al. (2008)
	Female		p	20	97	4.85	38.77	0.0135	0.0150	0.0806	Baumsteiger et al. (2008)
	Male	2003	q	110	493	4.48	62.03	0.0261	0.0303	0.118	Baumsteiger et al. (2008)
	Female		r	155	603	3.89	27.29	0.0319	0.0371	0.0715	Baumsteiger et al. (2008)

For each study, we include species and year, in addition to the known number of parents (n_{par}), number of juvenile offspring assigned to parents (n), sample mean ($\hat{\mu}$) and sample variance ($\hat{\nu}$) of reproductive success, and priors on the proportion of juvenile population sampled (f_{min} and f_{max}). ID refers to the plots labels in Figures 1 and 2. Type I error rate, estimated from the post hoc application of the model, refers to the probability of falsely rejecting the null hypothesis that offspring were distributed amongst parents according to the known true distribution within model simulations.

sex—these parents are assumed to produce a finite vector of offspring \mathbf{n}_{off} , where each element of \mathbf{n}_{off} represents the true (but unknown) discrete number of offspring per parent. Each study reports the number of offspring sampled n ($n \leq \sum \mathbf{n}_{\text{off}}$). The sample mean and variance of the offspring produced per parent can be represented by $\hat{\mu}$ and $\hat{\nu}$, respectively. These sample estimates along with n_{par} and n are known for each dataset. Our aim is to infer the true population mean and variance (μ, ν) from the sample size (n) and sample summary statistics (i.e., $\hat{\mu}$ and $\hat{\nu}$).

The reproductive success for each cohort was assumed to follow a negative binomial distribution, which is ideal when populations have skewed reproductive success, or a large fraction of the population fails to produce. Several alternative forms of the negative binomial distribution exist; in our analysis, we use the form familiar to ecologists, parameterized in terms of the mean and degree of overdispersion P ($P = \mu/\nu$); as P approaches 1, the negative binomial converges to the Poisson distribution (Hilborn and Mangel 1997). Initial model exploration with wide uniform priors (0.0, 1.0) on the negative binomial overdispersion parameter (P) suggested very low values for P posterior distributions, so we used (0.000001, 0.08) for all results presented here because this modification reduced computation time. We also assigned a uniform prior on the proportion of offspring sampled ($f = \hat{\mu}/\mu$), which enabled us to treat μ as a derived parameter.

Within each study, we estimated realistic upper and lower bounds on the total juvenile population size. These bounds were then used to create upper and lower limits for the priors on f for each cohort in the model, which were

calculated as the total number of sampled juveniles assigned parentage divided by the lower and upper bound of the estimated juvenile abundance. Baumsteiger et al. (2008) trapped downstream migrants and estimated the total population of juveniles using mark-recapture techniques; we used their 95% confidence limits to bound f . The other 3 studies did not provide an estimate of the number of fish in the total population at the life stage sampled. For these data, we bounded f based on an estimate of the total number of hatched juveniles (i.e., fry) and seaward migration smolts because all authors sampled juveniles between these 2 life stages. Ford et al. (2006) provided a numerical smolt estimate, and we used a species-specific fry-to-smolt survival (0.165; Quinn 2005, Table 15-1) to estimate the number of fry. For the studies by Garant et al. (2001) and Seamons et al. (2004), we used female counts, species-specific fecundity values (Garant et al. 2001; Quinn 2005), a relatively high egg-to-hatching survival rate (0.38; Quinn 2005, Table 15-1), and a relatively low value for egg-to-smolt survival (0.014; Quinn 2005, Table 15-1) to bound f . These last 2 survival rates were selected because they represented the maximum egg-to-hatching survival and the minimum egg-to-smolt survival within the average values reported by Quinn (2005) for Pacific salmon.

Estimating the unknown reproductive success parameters is challenging because it requires integrating over the potential vectors of offspring. We embedded the technique of Solow (1994) within a Bayesian Sampling Importance Resampling (SIR) algorithm (Rubin 1988). The SIR algorithm is similar to a weighted bootstrap, with weights of different parameter vectors being proportional to the

posterior probability of that vector given the data. The probability of the true population mean and variance given the reported sample statistics was

$$\Pr(\mu, P | \hat{\nu}, \hat{\mu}) = \int \Pr(\hat{\nu}, \hat{\mu} | n_{\text{off}}, \mu, P) f(n_{\text{off}}) dn_{\text{off}},$$

where n_{off} represents the vector of offspring produced per parent (and is unknowable). To integrate over n_{off} , we adopted the following procedure:

1. Given a proposed draw of model parameters (f^* , P^*), calculate the negative binomial mean μ^* as a derived parameter and use (μ^* , P^*) to generate a large number ($I = 300$) of hypothetical offspring populations (each produced by n_{par} parents). The total number of (f^* , P^*) draws varied by cohort because of differences in computation time but ranged from 5871 to 87 481 (median = 19 668).
2. From each of these populations of offspring, we repeatedly sampled without replacement a large number of times ($J = 300$) to generate hypothetical samples, each of size $\sum n_{\text{off}}$.
3. For each of the hypothetical offspring populations, we tabulated the number of times, T , that the mean and variance of the hypothetical sampled populations equaled the mean and variance in the observed data (i.e., $\mu^* = \hat{\mu}$ and $\nu^* = \hat{\nu}$). The posterior probability for this (f^* , P^*) pair is then calculated as $T/(J \times I)$.
4. After culling a small fraction (0.15%) of combinations with outlier values of $\mu > 1000$, we saved 100 draws in proportion to their posterior probabilities following the Hilborn-SIR procedure (Moore and Semmens 2008).

Of the 18 cohorts entered into the parameter estimation first stage of the model, all but one cohort produced realistic posterior estimates (the posterior distribution of cohort n was flat, indicating low information in the data). This provided 1700 (100 for each of 17 cohorts) reliable estimates of μ and P and essentially estimated ν as well because P was easily converted to ν ($\nu = \mu/P$).

Although we could not be certain of the reason why cohort n failed to produce informative posteriors, this cohort had the highest sample variance ($\hat{\nu}$) amongst the 4 cohorts measured by Ford et al. (2006). These 4 cohorts had the largest values of n_{par} by far (Table 1), and the 3 that did produce results took an extremely long computation time relative to the other 14 cohorts. We therefore speculate that the SIR algorithm struggled in parameter space with large values of both n_{par} and $\hat{\nu}$.

Stage Two: Offspring Assignment Simulations

The second stage of our model simulated the sampling and assignment process of offspring in order to identify sample size thresholds necessary to reach specific targets. For each of the 1700 μ and P combinations estimated above, we ran 250 simulations. Each simulation drew a random number from the negative binomial distribution based on μ and P to represent the reproductive success of each parent in the

population. Offspring were randomly sampled without replacement from this distribution one by one, and we evaluated the extent to which the samples assigned offspring represented the true patterns of reproductive success. This process was repeated for 4 different values of the number of parents in the population (n_{par}): 20, 50, 100, and 250. For the sake of simplicity, our procedures assumed no errors in the genetic assignments, as this issue has been addressed elsewhere (Morrissey and Wilson 2005; Kalinowski et al. 2007).

We used the χ^2 goodness-of-fit test to evaluate how well each sample of assigned offspring represented the true pattern of reproductive success in the simulated population. This was done by calculating type I and type II error rates for samples of 100–1000 assigned offspring in increments of 100 assigned offspring. To estimate the type I error rate, we compared the number of offspring assigned to each parent with the known relative reproductive success distribution specific to that simulation. The expectation was a failure to reject H_0 , and any simulations that rejected H_0 represented a type I error. To estimate our type II error rate, we compared the simulated offspring assignment with a uniform distribution where each parent produced an equal number of offspring. The expectation was to reject H_0 because offspring were assigned negative binomially, and a failure to reject would represent a type II error. For both types of errors, the χ^2 goodness-of-fit test used $\alpha = 0.05$.

Results

Posterior distributions of the negative binomial P parameter ($P = \mu/\nu$) were very small with a maximum value across all 1700 SIR runs of 0.044 (Figure 1). As values close to 0 indicate overdispersed data, this provided strong support for the negative binomial over the simpler Poisson distribution. Cohorts with wider priors on $f(a-i)$ tended to have wider posterior distributions of both P and μ than those with narrower priors on $f(k-r)$; Figures 1 and 2). Posterior distributions of μ tended to be left skewed for the Atlantic salmon, steelhead trout, and coho salmon cohorts (Figure 2a–m), and the median across all SIR retained values of μ was 111 juvenile offspring per parent. Across all cohorts, estimates of the population mean (μ , Figure 2) were substantially larger than values for the sample mean ($\hat{\mu}$, Table 1), indicating that researchers only sampled a small fraction of the total juvenile population.

As more samples were assigned to parents, both type I and type II errors decreased in a manner dependent on n_{par} , the number of single sex parents in the simulated population. At a given number of assigned offspring, both error types were greatest at the largest values of n_{par} (Figure 3). The type I error rate, or the probability of falsely rejecting the null hypothesis that offspring were distributed amongst parents according to the known true distribution, could be reduced below 5% by sampling 200 juvenile offspring from a cohort of 20 parents of a single sex, 400 offspring from a cohort of 50 parents, or 600 offspring from a cohort of

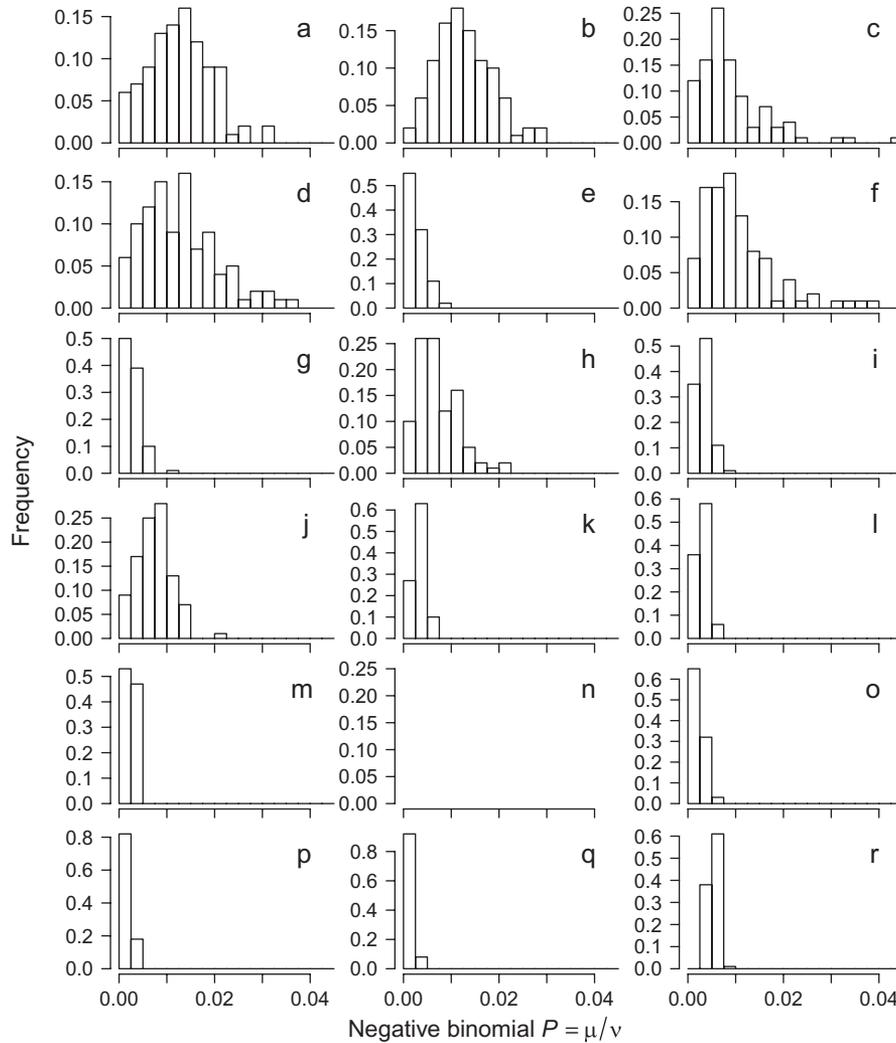


Figure 1. Posterior distributions of P (μ/v) for each cohort. Panels (a) and (b): Atlantic salmon, (c–j): steelhead trout (c and d: brood year 1997, e and f: 1998, g and h: 1999, and i and j: 2000), (k–n): coho salmon (k and l: 2000, m and n: 2001), and (o–r): Chinook salmon (o and p: 2002, q and r: 2003). Males are plotted in panels a, c, e, g, i, k, m, o, and q; females are plotted in panels b, d, f, h, j, l, n, p, and r. These panel labels match the ID column in Table 1.

100 parents (Figure 3a). For a cohort of 250 parents, sampling 1000 offspring brought the type I error rate below 10% but not 5% (Figure 3a). Type II error rates, or the probability of failing to reject a null hypothesis of uniform reproductive success, were all $<5\%$ except for simulations in which only 100 offspring were assigned parentage to a cohort of 250 parents (Figure 3b). Type II error therefore appeared to be of little worry in the salmonid fishes parameter space under consideration, so subsequent analysis focused on type I error rates.

In order to determine if the reproductive success values used to parameterize the assignment portion of the model had a significant effect on the type I error rate, we fit a quasibinomial generalized linear model (logit link) designed for overdispersed binomial data (Venables and Ripley 2002). The model was fit separately for each value of n_{par} and took the form:

$$\text{type I error rate} \sim P + \text{offspring} + P \times \text{offspring},$$

where offspring was the number of offspring assigned parentage and $P = (\mu/v)$. There was a negative relationship between P and type I error rate such that the errors were more likely for simulations parameterized with low values of P (Figure 4, Table 2).

In order to demonstrate the utility of our model to assess type I error rate post hoc, we reparameterized stage 2 of the model with observed values of n_{par} and n for each of the 100 μ and P combinations estimated from each cohort. One thousand simulations were run for each μ and P combination, and the resulting type I error rates ranged from 0.0189 to 0.239 (Table 1). The relationship between the simulated post hoc error rate and sample mean ($\hat{\mu} = n/n_{\text{par}}$) was exponential (Figure 5), and an ordinary least squares linear model of the form $\log(\text{error rate}) \sim \log(\hat{\mu})$ indicated

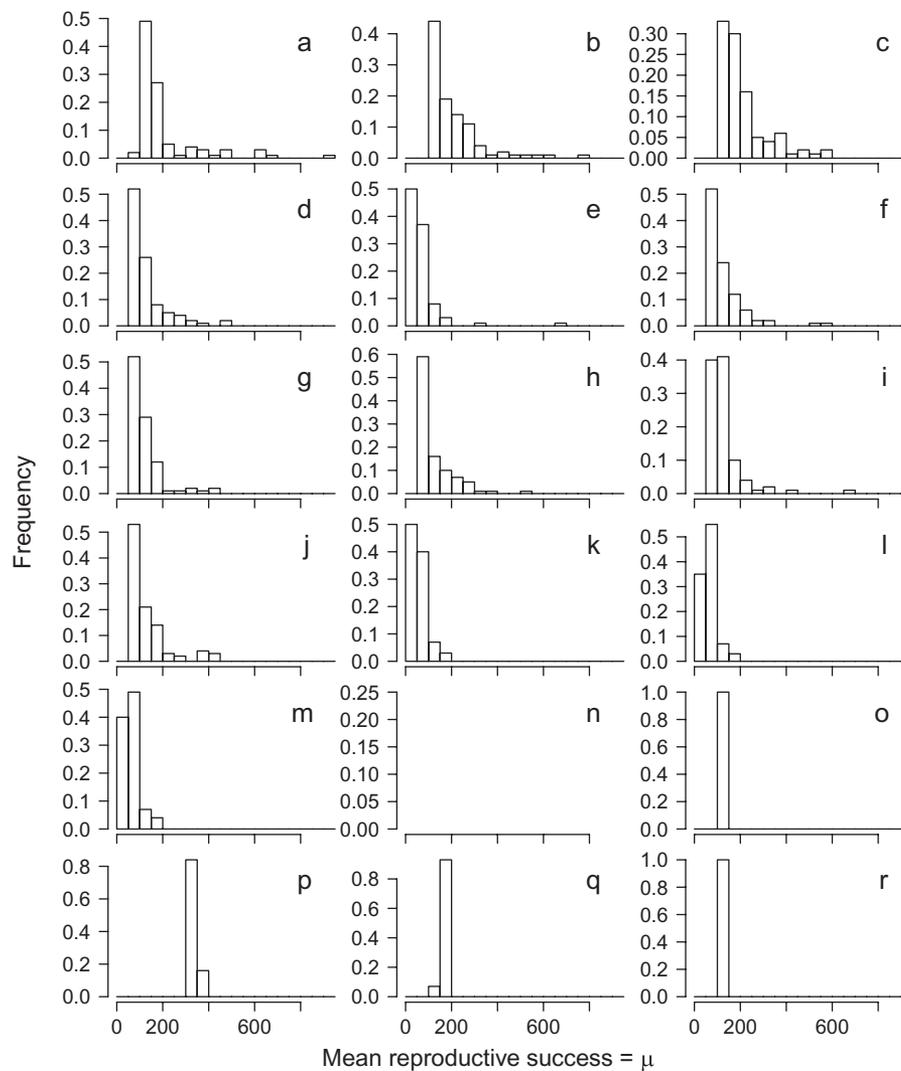


Figure 2. Posterior distributions of population mean reproductive success (μ) for each cohort. Panels (a) and (b): Atlantic salmon, (c–j): steelhead trout (c and d: brood year 1997, e and f: 1998, g and h: 1999, and i and j: 2000), (k–n): coho salmon (k and l: 2000, m and n: 2001), and (o–r): Chinook salmon (o and p: 2002, q and r: 2003). Males are plotted in panels a, c, e, g, i, k, m, o, and q; females are plotted in panels b, d, f, h, j, l, n, p, and r. Plot labels match the ID column in Table 1.

that $\hat{\mu}$ explained much of the variation in simulated error rate ($r^2 = 0.89$, $p < 0.0001$). This model suggested that studies with $\hat{\mu} > 6.2$ had type I error rates < 0.05 (Figure 5). In addition, type I error rates tended to be lower for cohorts with wide posterior estimates of the negative binomial P parameter (cohorts a–d and f: Figure 1) and for female cohorts within a given brood year (Table 1).

Discussion

In this paper, we developed a model to provide sampling guidance for reproductive success studies of taxa that produce more offspring than can feasibly be sampled. The most direct application of our results is to DNA-based

parentage investigations in fishes, amphibians, and insects in which the entire population of juvenile offspring may number in the thousands. We review the primary findings of our model and use them to provide practical guidance in designing reproductive success studies.

Our model indicated that for species in which reproductive success variance is often high, type I error but not type II error should be a concern for researchers quantifying the number of offspring per parent. Stage one of the model provided strong support for modeling salmonid reproductive success with the negative binomial distribution and indicated that offspring production was highly skewed amongst parents ($P < 0.05$, Figure 1) and not distributed randomly according to the Poisson distribution (in which case, P would have been much closer to one). Given this

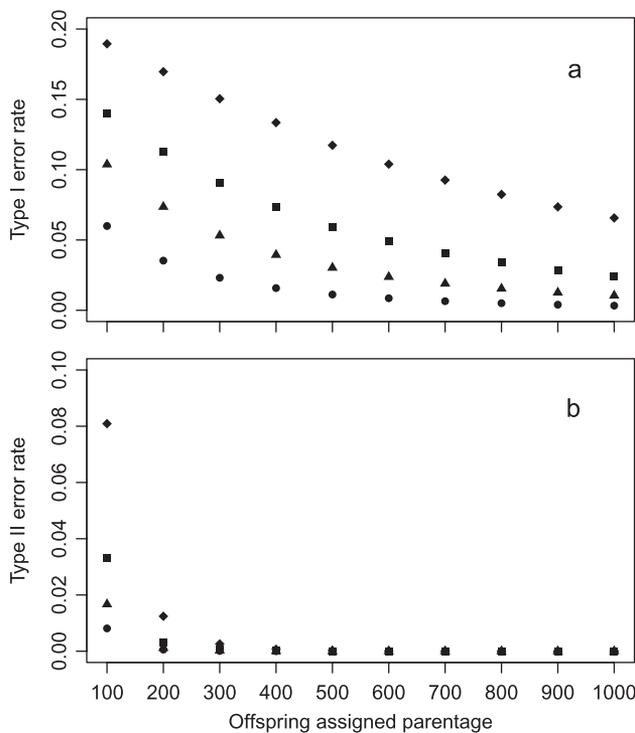


Figure 3. Relationship between (a) type I and (b) type II error rates and number of assigned offspring across all simulations. Different symbols represent different values of n_{par} : circles are 20, triangles are 50, squares are 100, and diamonds are 250.

result, it was not surprising that it was highly unlikely for a simulated sample to erroneously show uniform offspring production across parents. Therefore, statistical power (i.e., $\beta = 1$ —type II error rate) to reject a null hypothesis of uniform offspring production in reproductive success studies is likely to be high in systems similar to those on which our model is based. However, simply assigning parentage to enough offspring to reject a null hypothesis of uniform offspring production does not guarantee an accurate measurement of reproductive success. Our model indicated that sampled reproductive success distributions significantly deviated (χ^2 goodness-of-fit test p value < 0.05) from the known simulated reproductive success distribution at rates often > 0.05 and as high as 0.24, even when hundreds of offspring were assigned parentage (type I error rate: Figures 3a and 5).

Three key parameters, one of which is directly controlled by researchers, governed the type I error rate: the number of offspring assigned parentage, number of parents (n_{par}), and the variance in parent reproductive success (P). The most obvious observation is that researchers can reduce type I error rate by sampling more offspring and assigning them parentage. Although this goes without saying, our model provides quantitative guidelines for maximizing the trade-off between increased sampling effort and increased accuracy in estimating the reproductive success distribution

(Figures 3a and 5). As a general rule of thumb, our analysis of post hoc type I error rates suggests the ratio of assigned offspring to the number of parents (i.e., the sample mean, $\hat{\mu}$) is a good predictor of type I error rate (Figure 5). Sampling enough offspring so that this ratio is > 6.2 will reduce the error rate below 0.05. In addition, for a given number of offspring assigned parentage, reproductive success distributions were more accurately measured in cohorts with fewer adult breeders (n_{par}) and lower variance in reproductive success (i.e., larger value of negative binomial P). Although P may be unfamiliar to many behavioral ecologists, it is a simple ratio of 2 common metrics (mean and variance), and we found it to be a better predictor of type I error rate than other variance-based expressions (data not shown).

We argue that researchers should consider these 3 parameters when designing reproductive success studies and offer practical guidance for sampling. Although n_{par} is not directly under investigative control, researchers initiating reproductive success studies could choose systems in which previous adult abundance estimates indicate that research budgets permit the sampling, genotyping, and parentage assignment of sufficient numbers of juvenile offspring to reduce type I error rate to acceptable levels. However, sites for reproductive success research, at least in salmonids, are often selected opportunistically based on logistical factors such as the presence of a stream-spanning dam or weir at which all upriver migrating breeding adults (i.e., the parents) can be sampled (Seamons et al. 2004; Ford et al. 2006) or transplanting of adult salmon to previously unoccupied or underutilized habitats (Garant et al. 2001; Baumsteiger et al. 2008).

If a study system has already been selected and parents are numerous (e.g., > 100 per sex), researchers have 2 options. First, they may sample and assign parentage to many juvenile offspring, and the curves in Figures 3a and 5 provide explicit guidelines of the rate at which type I error is reduced with increasing numbers of assigned offspring. Second, researchers may opt to sample adult offspring in addition to or instead of juvenile offspring. If adult offspring can be sampled completely, assigning parentage to them will provide a very accurate measurement of reproductive success. For migratory salmonids that home to the natal site, this requires waiting until reproductive maturity for tissue collection (ca. 3–5 years for the species considered here), but complete sampling of adult offspring is often made possible by the same weir or dam structures used to collect tissues from the parents. In species that are not philopatric, exhaustive sampling of adult offspring will probably be difficult but could be achieved in experimental enclosures (e.g., Gopurenko et al. 2007) that prevent emigration or by investigating taxa with limited dispersal capability. It should be noted, however, that sampling adult rather than juvenile or embryonic offspring would provide less accurate measurements of other important reproductive parameters, such as mating success.

Although mean and variance of reproductive success cannot be known prior to sampling, anticipated results for these parameters can also inform sampling protocols. For

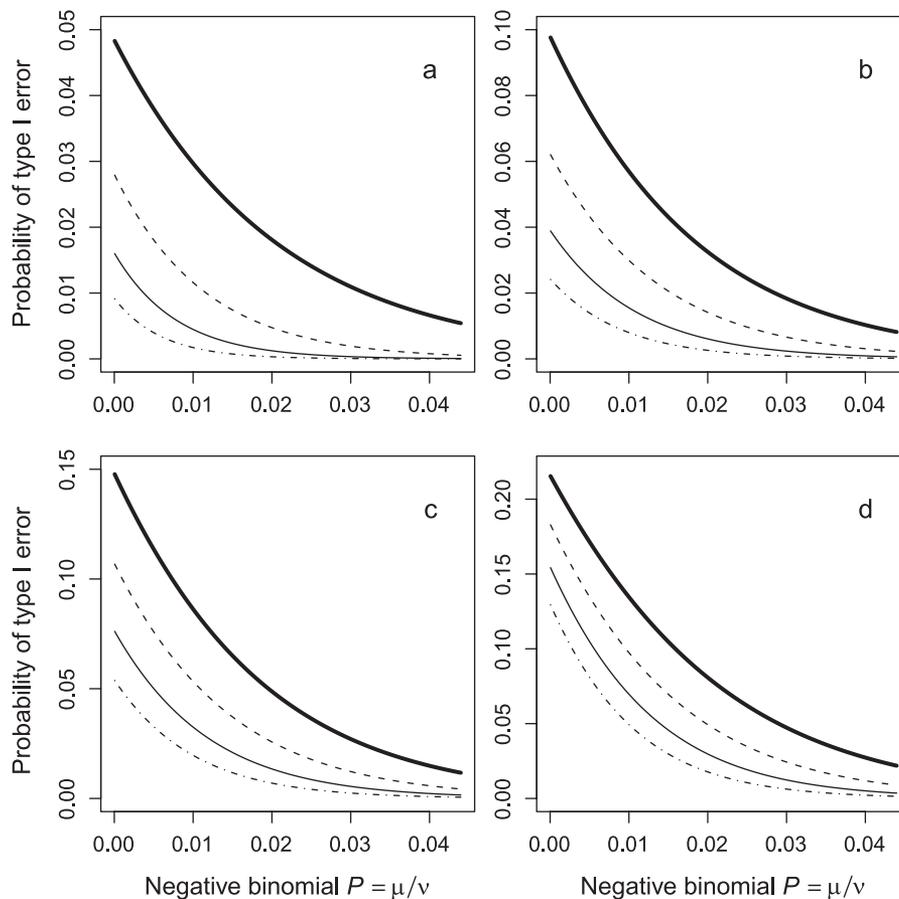


Figure 4. Probability of type I error predicted by quasibinomial generalized linear model for (a) $n_{\text{par}} = 20$, (b) $n_{\text{par}} = 50$, (c) $n_{\text{par}} = 100$, and (d) $n_{\text{par}} = 250$. The form of the model was type I error $\sim P + \text{offspring} + P \times \text{offspring}$ where $P = \mu/v$ and offspring was the number of offspring assigned parentage. Offspring = 200 for the thick solid line, offspring = 400 for the dashed line, offspring = 600 for the thin solid line, and offspring = 800 for the dot-dash line.

example, salmonid fishes have a polygynandrous mating system in which males compete for access to females (Esteve 2005), and in these systems, males will typically have a higher variance in reproductive success than females (Shuster and Wade 2003). Combined with the result that type I errors were more common in simulations with lower values of P , this indicates that researchers will usually have

to assign parentage to more offspring to accurately represent male reproductive success compared with female reproductive success. This conclusion was supported by the post hoc error rates calculated for each of 17 cohorts in the analysis: Male cohorts had a higher estimated type I error rate than female cohorts from the same brood year (Table 1).

Table 2 Results of quasibinomial generalized linear model: type I error $\sim P + \text{offspring} + P \times \text{offspring}$ where $P = \mu/v$ and offspring was the number of offspring assigned parentage

n_{par}	Parameter estimate				Deviance	
	Intercept	P	Offspring	$P \times \text{offspring}$	Residual	Null
20	-2.41	-11.6	-0.00284	-0.196	43 675	108 104
50	-1.73	-40.6	-0.00246	-0.0903	60 094	163 098
100	-1.38	-46.5	-0.00186	-0.0728	59 847	179 845
250	-1.09	-40.9	-0.00102	-0.0809	48 900	170 445

The model used a logit link and was fit separately for each value of n_{par} . All parameter estimates were significantly different from zero (p value < 0.0001).

Knowledge of reproductive success mean (μ) and variance (v) would also permit application of our model to taxa other than the salmonid fish data presented here. Our model is generic with respect to species: All that it requires are the input parameters listed in Table 1. It would be relatively straightforward to collate these input data for additional species under consideration for a new parentage study and then run the model to estimate the number of offspring that must be sampled to accurately represent reproductive success. Such an approach would help determine if the results presented here are exclusive to species with highly skewed reproductive success or general regardless of reproductive success variance. To facilitate the use of our model, all code is provided as open source at <http://www.ecologybox.org> (project name: Offspring sampling in reproductive success studies, last accessed June 7, 2011).

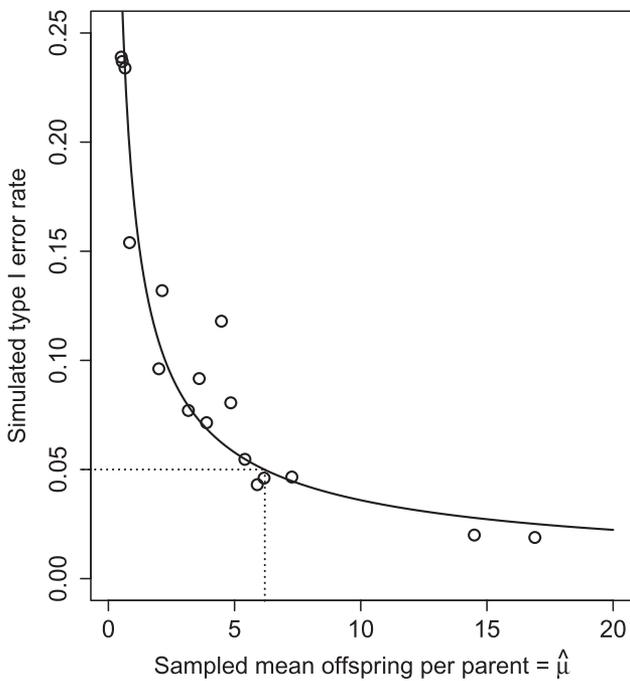


Figure 5. Relationship between simulated post hoc type I error rate and estimated mean reproductive success. Points represent cohorts used in this analysis (Table 1), and the predicted values (line) are estimated an ordinary least squares regression where both predictor and response variables were log transformed ($p < 0.0001$, $r^2 = 0.89$). The type I error rate becomes less than 5% when the number of offspring sampled per parent is >6.2 .

The model displayed 2 related limitations that suggest our approach is best suited for smaller populations (i.e., ca. <300 parents per sex). First, as previously discussed by Solow (1994), the SIR algorithm is extremely slow because it requires numerically integrating over potential vectors of offspring for each proposed parameter combination. We found this to be particularly true for cohorts k–m, which had the greatest number of parents and took several days to run. Second, the failure of cohort n to produce reliable posteriors indicates that the SIR algorithm may perform poorly in populations with several hundred parents. The primary consequence of this result in our study was a reduced sample size for the second stage of the model. However, this could be more problematic for species in which reproductive success mean and variance metrics are rare if algorithm failure exacerbates a lack of input data.

Two other aspects of sampling deserve consideration in addition to the numbers suggested here, namely the spatial distribution of sampling and the success rate of parentage assignments. First, sampling should cover the spatial extent of the juvenile population, which might often be estimated by the spatial distribution of breeding adults. In the wild, each female's offspring are most likely located in the vicinity of her nest site, and the probability of collecting her offspring will decrease as juvenile samples are captured

further and further away. Sampling only a portion of habitats in which juveniles are found would bias reproductive success estimates by inflating the productivity of females that bred in or near the sampled areas. Secondly, our model provides estimates of the required number of assigned offspring, and this will probably be less than the total number of samples collected in the field. There exist a variety of methods to assign offspring (reviewed by Jones and Ardren 2003; Jones et al. 2010), and these will vary in the assignment success rate. Failure to assign offspring could result from not sampling enough parents (Marshall et al. 1998) or not collecting enough genetic data, both in terms of the number of loci and the polymorphism of the markers used (Bernatchez and Duchesne 2000). By considering these factors, researchers can estimate the number of unassigned offspring, and then adjust the number of samples collected in the field upward from the sample sizes suggested here.

Finally, the goal of most empirical reproductive success studies is not only to quantify reproductive success but to relate reproductive success to individual quantitative traits, behavioral phenotypes, or population of origin. Our results have indicated that the insufficient offspring sampling could lead to inaccurate estimates of reproductive success, but it was beyond the scope of our model to explicitly predict how these errors could bias statistical models of the relationship between specific traits and reproductive success. However, it seems likely that inadequate offspring sampling would introduce observation error into such models, which could undermine the power of statistical conclusions regarding the influence of individual phenotypes on reproductive success. Therefore, the consequence of sampling too few offspring could be type II error of a different sort than that presented in this paper in which researchers fail to detect a biologically important statistical relationship between the trait of interest and reproductive success. Indeed, we suggest that the lack of knowledge regarding the effects of methodological errors on the ultimate analytical conclusions in parentage or pedigree studies is a general problem that merits future research.

Funding

J.H.A. and S.M.C. were supported by the H. Mason Keeler Endowment, administered by the School of Aquatic and Fishery Sciences at the University of Washington. J.H.A. was also supported in part by a grant from Washington Sea Grant, University of Washington, pursuant to National Oceanic and Atmospheric Administration Award No. NA04OAR4170032 and NA07OAR4170007, Project No. R/F-148 and R/F-159. E.W.'s contribution was funded by a fellowship from the National Research Council.

Acknowledgments

We thank Andre Punt for guidance in constructing our model, Todd Seamons for thoughtful discussions on parentage assignment, and 2

anonymous reviewers for helpful comments on an earlier draft of the manuscript.

References

- Avise JC, Jones AG, Walker D, DeWoody JA. 2002. Genetic mating systems and reproductive natural histories of fishes: lessons for ecology and evolution. *Annu Rev Genet.* 36:19–45.
- Baumsteiger J, Hand DM, Olson DE, Spateholts R, FitzGerald G, Ardren WR. 2008. Use of parentage analysis to determine reproductive success of hatchery-origin spring Chinook salmon outplanted into Shitike Creek, Oregon. *N Am J Fish Manag.* 28:1472–1485.
- Bernatchez L, Duchesne P. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can J Fish Aquat Sci.* 57:1–12.
- Clutton-Brock TH. 1988. *Reproductive success: studies of individual variation in contrasting breeding systems.* Chicago (IL): The University of Chicago Press.
- DeWoody JA, DeWoody YD, Fiumera AC, Avise JC. 2000. On the number of reproductives contributing to a half-sib progeny array. *Genet Res.* 75:95–105.
- Esteve M. 2005. Observations of spawning behavior in Salmonidae: *Salmo, Oncorhynchus* and *Salvelinus*. *Rev Fish Biol Fish.* 15:1–21.
- Ford MJ, Fuss H, Boelts B, LaHood E, Hard J, Miller J. 2006. Changes in run timing and natural smolt production in a naturally spawning coho salmon (*Oncorhynchus kisutch*) population after 60 years of intensive hatchery production. *Can J Fish Aquat Sci.* 63:2343–2355.
- Garant D, Dodson JJ, Bernatchez L. 2001. A genetic evaluation of mating system and determinants of individual reproductive success in Atlantic salmon (*Salmo salar* L.). *J Hered.* 92:137–145.
- Garant D, Dodson JJ, Bernatchez L. 2003. Differential reproductive success and heritability of alternative reproductive tactics in wild Atlantic salmon (*Salmo salar* L.). *Evolution.* 57:1111–1141.
- Gopurenko D, Williams RN, DeWoody JA. 2007. Reproductive and mating success in the small-mouthed salamander (*Ambystoma texanum*) estimated via microsatellite parentage analysis. *Evol Biol.* 34:130–139.
- Gopurenko D, Williams RN, McCormick CR, DeWoody JA. 2006. Insights into the mating habits of the tiger salamander (*Ambystoma tigrinum tigrinum*) as revealed by genetics parentage analyses. *Mol Ecol.* 15:1917–1928.
- Hilborn R, Mangel M. 1997. *The ecological detective.* Princeton (NJ): Princeton University Press.
- Hinrichsen RA. 2003. The power of experiments for estimating relative reproductive success of hatchery-born spawners. *Can J Fish Aquat Sci.* 60:864–872.
- Jones AG, Ardren WR. 2003. Methods of parentage analysis in natural populations. *Mol Ecol.* 12:2511–2523.
- Jones AG, Arguello JR, Arnold SJ. 2002. Validation of Bateman's principles: a genetic study of sexual selection and mating patterns in the rough-skinned newt. *Proc R Soc Ser B Biol Sci.* 269:2533–2539.
- Jones AG, Small CM, Paczolt KA, Ratterman NL. 2010. A practical guide to methods of parentage analysis. *Mol Ecol Res.* 10:6–30.
- Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol.* 16:1099–1106.
- Liebig EB, Cabe PR, Jaeger RG, Leberg PL. 2006. Multiple paternity in a salamander with socially monogamous behavior. *Mol Ecol.* 15:4153–4160.
- Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol.* 7:639–655.
- Moore JW, Semmens BX. 2008. Incorporating uncertainty and prior information into stable isotope mixing models. *Ecol Lett.* 11:470–480.
- Morrissey MB, Wilson AJ. 2005. The potential costs of accounting for genotypic errors in molecular parentage analyses. *Mol Ecol.* 14:4111–4121.
- Neff BD, Repka J, Gross MR. 2000. Statistical confidence in parentage analysis with incomplete sampling: how many loci and offspring are needed? *Mol Ecol.* 9:529–539.
- Quinn TP. 2005. *The behavior and ecology of Pacific salmon and trout.* Seattle (WA): University of Washington Press.
- Rubin DB. 1988. Using the SIR algorithm to simulate posterior distributions. In: Bernardo JM, DeGroot MH, Lindley DV, Smith AFM, editors. *Bayesian statistics 3.* New York: Oxford University Press. p. 395–402.
- Seamons TR, Bentzen P, Quinn TP. 2004. The effects of adult length and arrival date on individual reproductive success in wild steelhead trout (*Oncorhynchus mykiss*). *Can J Fish Aquat Sci.* 61:193–204.
- Seamons TR, Bentzen P, Quinn TP. 2007. DNA parentage analysis reveals inter-annual variation in selection: results from 19 consecutive brood years in steelhead trout. *Evol Ecol Res.* 9:409–431.
- Shuster SM, Wade MJ. 2003. *Mating systems and strategies.* Princeton (NJ): Princeton University Press.
- Solow AR. 1994. On the Bayesian estimation of the number of species in a community. *Ecology.* 75:2139–2142.
- Tatarenkov A, Healey CIM, Grether GF, Avise JC. 2008. Pronounced reproductive skew in a natural population of green swordtails, *Xiphophorus helleri*. *Mol Ecol.* 17:4522–4534.
- Venables WN, Ripley BD. 2002. *Modern applied statistics with S.* New York: Springer.
- Zar JH. 1999. *Biostatistical analysis.* Upper Saddle River (NJ): Prentice-Hall.

Received January 11, 2011; Revised May 18, 2011
Accepted May 18, 2011

Corresponding Editor: Jose Lopez