

# Indirect genetic estimates of breeding population size in the polyploid green sturgeon (*Acipenser medirostris*)

J. A. ISRAEL and B. MAY

Genomic Variation Laboratory, Department of Animal Science, 1 Shields Avenue, University of California, Davis, CA 95616, USA

## Abstract

The utility of genetic measures for kinship reconstruction in polysomic species is not well evaluated. We developed a framework to test hypotheses about estimating breeding population size indirectly from collections of outmigrating green sturgeon juveniles. We evaluated a polysomic dataset, in allelic frequency and phenotypic formats, from green sturgeon to describe the relationship among known progeny from experimental families. The distributions of relatedness values for kin classes were used for reconstructing green sturgeon pedigrees from juveniles of unknown relationship. We compared three rarefaction functions that described the relationship between the number of kin groups and number of samples in a pedigree to estimate the annual abundance of spawners contributing to the threatened green sturgeon Southern Distinct Population Segment in the upper Sacramento River. Results suggested the estimated abundance of breeding green sturgeon remained roughly constant in the upper Sacramento River over a 5-year period, ranging from 10 to 28 individuals depending on the year and rarefaction method. These results demonstrate an empirical understanding for the distribution of relatedness values among individuals is a benefit for assessing pedigree reconstruction methods and identifying misclassification rates. Monitoring of rare species using these indirect methods is feasible and can provide insight into breeding and ontogenetic behaviour. While this framework was developed for specific application to studying fish populations in a riverscape, the framework could be advanced to improve genetic estimation of breeding population size and to identify important breeding habitats of rare species when combined with finer-scaled sampling of offspring.

*Keywords:* gene trees, kinship, polyploidy, population estimation

*Received 1 October 2009; revision received 15 December 2009; accepted 28 December 2009*

## Introduction

Knowledge of the genetic relationships among individuals, when combined with additional field information, can provide insight into influential physical and biological processes affecting these patterns. Information about dispersal and emigration rates is available when the recapture of genotypes is evaluated over space, while when these same genotypes are considered over time, population parameters like survival, state transition

rates, and the finite rate of change can be determined. Analytical approaches for determining the relationship among individuals have expanded the potential for the evaluation of pedigrees when parental information or samples are not available (Almudevar & Field 1999; Smith *et al.* 2001; Beyer & May 2003; Konovalov *et al.* 2004; Wang 2004a). These approaches have been used to estimate population size (Herbinger *et al.* 2006) and examine reproductive behaviours (Gottelli *et al.* 2007; Makinen *et al.* 2007).

Pedigree studies provide information about contemporary patterns of genetic relationships within cohorts giving insight into population demographics (Konovalov

Correspondence: Joshua A. Israel, Fax: +1 530 754 0175; E-mail: jaisrael@ucdavis.edu

*et al.* 2004; Wang 2004a). However, these studies have focused on diploid data where interpreting genetic data and applying the statistical models is relatively simple due to the well-understood laws of Mendelian inheritance. Molecular datasets of polyploid fish and plant species are increasingly common, and fine scale molecular studies on population size estimation, dispersal and invasion, and spatial organization of genetic diversity are becoming more common (Vekemans & Hardy 2004; Guillot *et al.* 2009). Green sturgeon (*Acipenser medirostris*) are functional tetraploids (Ludwig *et al.* 2001) and the majority of their microsatellite loci show four gene doses; thus, alternative analytical methods are necessary for using genetic relatedness data to reconstruct pedigrees and partition groups of similar kinship in this species. These green sturgeon data can be formatted using the observed tetrasomic banding pattern with four gene doses or as allelic phenotypes, where bands are scored as present/absent. An analysis of inter-individual relatedness data can provide information into the population ecology of green sturgeon. A comparison of these data formats demonstrate the utility of polysomic datasets for fine scale molecular studies relevant to understanding the ecology of species of interest.

In rivers, green sturgeon juveniles are presumably the most abundant life history stage and adults are highly migratory, uncommon, and difficult to sample. The production of juvenile white sturgeon (*A. transmontanus*), another sturgeon species that spawns in the Sacramento River, has been shown to be positively correlated with greater river flows between April and July (Kohlhorst *et al.* 1991). We hypothesized a similar relationship and that estimates of green sturgeon breeding population size would be positively correlated with river flow. During this study, flows on the Sacramento River were variable and demonstrated four of five water year types (Table 1), and we predicted a greater number of kin

groups, representing the contribution of more spawners, would be observed during the wetter years in the reach of the Sacramento River that was surveyed. This information was considered in evaluating reproductive and demographic factors, which may influence breeding success.

The reconstruction of pedigrees and partitioning of individuals based on relatedness measures provides a way to assess the kin structure and demographics of populations, and a variety of estimators and genetic methods have been developed for evaluating genealogical relationships (Blouin 2003). Although information about the quantitative relationship between genealogical classes has been explored through simulations by some programs (Goodnight & Queller 1999), the utility of information from known family pedigrees for partitioning cohorts of unknown juveniles into groups of similar kinship has not been well examined in the literature. If these data are effective for describing distinct kin relationships, then combining them with 'capture-recapture' methods can allow for estimating the number of breeders contributing to a collection of the wild cohort being examined. This approach uses rarefaction equations (Gotelli & Colwell 2001; Frantz *et al.* 2006) to estimate the abundance of kin groups in a collection, where individuals are randomized and iteratively added to reduce assumptions about the probability of encountering the same kin group. The ability to reliably monitor breeding populations is essential in conservation science, and if the relatedness among individuals within kin groups is quantified, it is possible to evaluate the 'recapture' of kin groups in the wild to yield information about an organisms' breeding population.

Estimating the population abundance of highly mobile species is difficult, yet managers and ecologists are increasingly interested in spatially and temporally fine-scaled demographic patterns of populations.

**Table 1** Collection dates, number of samples in collection, and size range for green sturgeon experimental families and wild samples collected at rkm 388.8 on the Sacramento River

Collection	Water year type	Number of samples	TL range (mm)	Mother	Father
Wild fry sampling dates					
June–July, 2002	Dry	14	26–52	Unknown	Unknown
June–July, 2003	Above normal	43	23–30	Unknown	Unknown
May–July, 2004	Below normal	22	24–41	Unknown	Unknown
June–July, 2005	Above normal	73	25–41	Unknown	Unknown
June–July, 2006	Wet	75	25–42	Unknown	Unknown
Experimental Families					
2002A		51	20–30	GS02–1	GS02–2
2002E		54	20–30	GS02–5	DOM0461
2003G		70	20–30	GS03–1	GS03–02
2003H		60	20–30	GS03–1	GS03–03

Genetic 'capture-recapture' methods using individual genotypes have been developed to estimate population size, as well as a variety of useful biological parameters (Pearse *et al.* 2001; Lukacs & Burnham 2005). These methods rely upon the repeated sampling of genotypes across space or time, and different information can be derived depending on the scale being evaluated. Collecting the necessary samples from highly migratory species can be challenging for such census studies, although it is frequently possible to gather tissue samples from more sedentary early life stages that reflect the genetic diversity of the breeding population. Genetic analyses of these young individuals can provide knowledge about the abundance, mating behaviour, and spawning habitat preferences of the breeding segment of the adult population.

This study characterized the distributions of relatedness values within experimental families with known pedigrees and utilized a kin group capture-recapture method for estimating the number of breeding green sturgeon in the wild. First, microsatellite genotypes of loci with observed tetrasomic and disomic banding patterns from multiple years of experimental families were used to characterize the distribution of relatedness and kinship coefficients among full-sibling, half-sibling, and unrelated pairs of individuals. The classification rates for these estimators of relatedness were evaluated to determine which was best for assessing the relationship among wild juvenile green sturgeon captured downstream of spawning grounds. Second, kinship data from five annual collections of wild green sturgeon fry from the Sacramento River was used to reconstruct kin group clusters by partitioning individuals who were full siblings. The relationship between the number of samples being examined and the number of full sibling groups observed was modelled to estimate the number of full sibling groups using a set of accumulation functions. The potential assumptions associated with different mating systems on the estimated number of full-sibling kin groups are discussed. The probability of encountering full siblings, by identifying members from the same family during different sampling dates within the annual collections, is evaluated as a basis for improving non-invasive genetic censuring of populations and developing future sampling protocols.

## Materials and methods

### Sample collection

Wild green sturgeon fry were collected at six rotary screw traps operated by the U.S. Fish and Wildlife Service in the Sacramento River below the gates of Red Bluff Diversion Dam, approximately 374 river kilome-

tres (rkms) upstream of the outlet of San Francisco Bay and approximately 96 rkm downstream of the presumed upper extent of green sturgeon spawning at Keswick Dam. These rotary screw traps were operated throughout the outmigration period of green sturgeon fry capturing fish encompassing the entire emigration period. Between 2002 and 2004, only green sturgeon incidentally killed during trap operations were collected. After 2005, all incidental mortalities were collected, as well as up to five additional, randomly-selected green sturgeon fry collected as genetic voucher samples daily. Table 1 summarizes the collections of green sturgeon from the Sacramento River. The taxonomic identity of wild samples was verified using a green sturgeon-specific mtDNA restriction fragment length polymorphism (Israel *et al.* 2004), since white sturgeon are morphologically similar and sympatric with green sturgeon in the Sacramento River.

Experimental families were created in 2002 and 2003 at the University of California, Davis as part of a larger study examining the biological characteristics of green sturgeon. Broodstock for these experimental families were collected and transported from the Klamath River (Northern Distinct Population Segment (DPS)), then spawned at UC Davis (Van Eenennaam *et al.* 2001, 2005). The two families from 2002 were each single pair matings, while the two 2003 families were spawned from the same mother with two sires (Table 1). All tissue samples were preserved in 95% ethanol, stored at room temperature, and DNA was isolated with the Wizard SV DNA extraction kit (Promega Corporation, Madison, WI, USA).

### Genetic data

Ten microsatellite DNA markers were amplified to yield genotype data (Israel *et al.* 2009). PCR conditions and genotyping dilutions were optimized for electrophoresis on the Base Station DNA Fragment Analyser (Bio-Rad). Briefly, a 0.5  $\mu\text{M}$  forward primer end-labelled with one of three PRISM fluorophores: NED, VIC, 6-FAM (Applied Biosystems Inc.) was used with each primer set. One  $\mu\text{L}$  of each diluted PCR product for each of the loci was combined with 2  $\mu\text{L}$   $\text{H}_2\text{O}$ , 1.45  $\mu\text{L}$  100% de-ionized formamide (Sigma), 0.5  $\mu\text{L}$  blue dextran loading dye, and 0.05  $\mu\text{L}$  ABI GeneScan-400 Rox internal size standard. PCR product/size standard mixtures were separated using 5.5% denaturing polyacrylamide gels. PCR products were analysed and genotypic data (based on DNA fragment size and gene dosage) were generated using Cartographer (version 1.2.6 g) DNA fragment analysis software. Two individuals were run on each gel to validate inter-gel allele size accuracy between gels.

Eight of ten microsatellite loci showed banding patterns with four gene doses, likely reflecting the tetraploid genome of green sturgeon. The number of bands was interpreted by relative peak intensity from the genotyping electropherogram to score four gene doses for these eight microsatellites assuming no null alleles. If three peaks were present, the largest peak was assigned two doses. If two peaks were present and equal in size, both peaks were scored as representing two doses. When two peaks were present but not equal in size, the largest peak was assigned three doses for that allele size. A single electrophoretic peak or four peaks were assigned as either four doses of a single allele or one dose of each of four alleles, respectively. Alternatively, to score loci in a presence/absence format, genotypic data were transformed into 'allele phenotypes' (Becher *et al.* 2000). Each phenotype scored alleles as present or absent, regardless of dosage. This presence/absence matrix combined all of the alleles at the ten loci identifying 137 markers for use in pedigree reconstruction analyses.

#### *Estimates of relatedness and classification rates with experimental families*

Allelic frequency and presence/absence datasets were formatted with TRANSFORMER (Caujape-Castells & Baccarani-Rosas 2005) for analysis in the computer program SPAGeDi (Hardy & Vekemans 2002). Two separate sets of collection data were formatted and used in SPAGeDi. The first consisted of the 235 individuals in the four Klamath River broodstock experimental families, and the second set contained the wild Sacramento River fry. The dataset of experimental families was used to evaluate eight estimators of relatedness ( $r$ ). Individuals from only one 2003 family and the two 2002 families were used to approximate the distribution of unrelated dyad values. Relationship coefficients for the four gene dose allelic frequency dataset proposed by Loiselle *et al.* (1995), Ritland (1996), Streiff *et al.* (1998) and Hardy & Vekemans (1999) were calculated in the computer program SPAGeDi. This allelic frequency dataset contained genotypes from the eight loci demonstrating four gene dose banding patterns. Kinship ( $F_{ij}$ ) and relatedness coefficients proposed by Hardy (2003) for allelic phenotypes were calculated with SPAGeDi. These estimators did not assume Hardy–Weinberg equilibrium in the collection, and thus required independent estimates of the population's inbreeding coefficient,  $F_{IS}$ . Independent  $F_{IS}$  values were estimated with AUTOTET (Thrall & Young 2000) for green collections of potential parents with the eight loci containing four gene doses. Lynch & Milligan's (1994) relatedness coefficient was calculated using the computer program RELATE (Rodzen *et al.* 2004)

with the allele phenotype dataset. For further analysis of classification rates and pedigree reconstruction, each pairwise relatedness values ( $r_i$ ) from the experimental families were recorded as  $1 - r$ . Thus, unrelated individuals appeared further apart than individuals of shared ancestry in dendrograms.

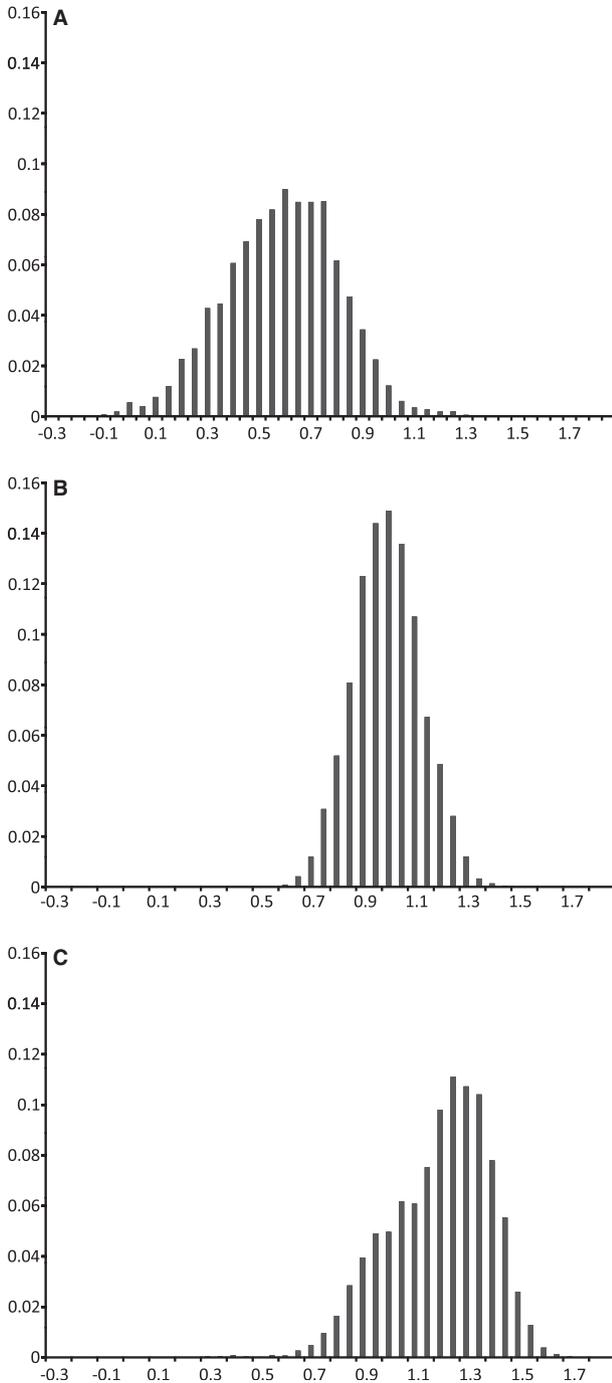
In an effort to identify the best estimator of relatedness for our purposes, the method from Blouin *et al.* (1996) was used to compare distributions of pairwise values from experimental families. We assessed type I and type II error rates for the relatedness estimators based upon the distributions of pairwise values calculated for dyads of known relationship classes (full sibling, half sibling, and unrelated). More specifically, error rates were determined by using the midpoint ( $r_{\text{midpoint}}$ ) between the means of the pairwise value distributions. The type I error rate was the proportion of values belonging to experimental family dyads that was greater than the midpoint value, while the proportion of pairwise values between individuals of the other relationship type being examined that was less than the cut-off relatedness value was the type II error rate (Fig. 1).

#### *Pedigree reconstruction of wild juveniles and partitioning kin groups*

The second dataset created from SPAGeDi (Hardy & Vekemans 2002) contained the 229 Sacramento River green sturgeon fry captured between 2002 and 2006). SPAGeDi estimated Hardy's (2003) relatedness coefficient relative to the pool of samples included in the dataset. A pairwise matrix, consisting of relatedness values for specific annual subsets of each dataset, was analysed with the unweighted pair group method using the arithmetic means (UPGMA) algorithm to construct a dendrogram based with the computer package PHYLIP's NEIGHBOR program (Felsenstein 2005). Each dendrogram was visualized with Tree Explorer in the computer package MEGA (Kumar *et al.* 2004) representing the samples in the dataset. Each dendrogram was visually inspected and a line drawn at  $r_{\text{midpoint}}/2$  was considered the partition relatedness value ( $r_{\text{partition}}$ ) for which individuals joining at a node of lesser value belonged to the same full sibling kin group. The estimated number of full sibling kin groups was calculated based on a dendrogram that included all the annual samples from one year.

#### *Estimating local breeder abundance*

The matrix of Hardy's (2003) relatedness coefficient values for all unknown fry was dissected into smaller matrices containing only the annual collections of



**Fig. 1** The distribution of  $1-F_{ij}$  values for dyads between full siblings (Panel A), half siblings (Panel B), and unrelated individuals (Panel C) calculated from controlled crosses using Hardy's (2003) relatedness coefficient values.

samples. Pairwise relatedness values of green sturgeon fry samples were added into the annual matrices randomly in a stepwise fashion without replacement. A UPGMA dendrogram was constructed following each addition of one new individual to the distance matrix.

The ability to accurately partition individuals in full sibling relationships allowed for a capture-recapture approach, which used kin groups, as the unit of resampling. The probability ( $p^*$ ) that a kin group is encountered at least once was calculated after all individuals were placed into the dendrogram.

SIGMAPLOT 9.0 (Systat Software, San Jose, CA, USA) was used to perform iterative nonlinear regression for assessing the asymptote of the line described by the relationship between the number of individuals in a UPGMA dendrogram ( $x$ ) and the number of clusters discovered in the UPGMA dendrogram ( $y$ ). Three different functions (Kohn *et al.* 1999; Valiere 2002; Eggert *et al.* 2003) were used for fitting an accumulation curve and deriving the maximum estimated number of kin groups. Ninety-five percent confidence intervals were constructed for the regression lines and expressed as  $\pm (1.96 * \text{Standard Error})$ . California Department of Water Resources water year types and daily flow data from rkm 414 (Bend Bridge) between March 1 and May 31 was averaged for the period. The observed number of kin groups annually was plotted against the average daily flow data using SIGMAPLOT 9.0 to evaluate the correlation between these two variables.

## Results

### *Relatedness and kinship coefficients*

To calculate relatedness and kinship estimators we examined 137 alleles in the allele phenotype dataset that included 40 alleles distinct to the annual wild datasets, 27 distinct to the experimental families, and 70 that were present in both sets of collections. Four experimental families contained a total of 27 495 dyads, of which 6891 were full-siblings, 4200 were half-siblings, and 16 404 were 'unrelated' dyads. Only 10 104 dyads between the two 2002 families and the 2003G family were included in the distribution of 'unrelated' values to minimize the potential influence of 2003 half-sibling on this distribution (Fig. 1). For the Klamath River experimental families, a  $F_{IS}$  value of 0.189 was calculated for adult green sturgeon collected over 3 years in the Lower Klamath River ( $n = 125$ ). The mean inbreeding coefficient for adult green sturgeon captured in San Pablo Bay in 2001 and 2004 ( $n = 219$ ) was 0.197, compared to  $F_{IS}$  values between 0.179 and 0.217 for the five juvenile Sacramento River green sturgeon collections. Theoretical relatedness coefficients in progeny of different relationships (Hedrick 2000) and means and standard deviations of seven relatedness estimators were compared (Table 2). All coefficients showed relative imprecision with moderate and high standard deviations, although the Hardy (2003) and Lynch & Milligan

**Table 2** Mean relatedness coefficients (standard deviations) for different classes of relatedness with seven algorithms for distributions of relationship estimators for juvenile green sturgeon from controlled matings described in Table 1

	Full sibling dyads	Half sibling dyads	Unrelated dyads
Expected R	0.50	0.25	0.00
Relatedness (Hardy 2003)	0.436 (0.220)	0.035 (0.130)	-0.170(0.194)
Kinship (Hardy 2003)	0.230 (0.131)	0.021 (0.077)	-0.101 (0.115)
Relatedness (Lynch and Milligan 1994)	0.491 (0.189)	0.328 (0.167)	0.083 (0.178)
Ritland (1996)	0.086 (0.057)	-0.007 (0.023)	-0.03 (0.025)
Loiselle <i>et al.</i> (1995)	0.147 (0.084)	0.003 (0.067)	-0.039 (0.055)
Streiff <i>et al.</i> (1998)	0.271 (0.216)	0.058 (0.19)	-0.134 (0.29)
Hardy and Vekemans (1999)	0.402 (0.23)	0.009 (0.152)	-0.106 (0.183)

(1994) relatedness estimators appeared to most closely approximate expected values for the allelic phenotype dataset and Hardy & Vekemans (1999) most closely approximated values for the allelic frequency dataset. The three allelic phenotype relatedness estimators were less biased for full siblings than estimators calculated with allelic frequency data, and overall the kinship estimators calculated with presence/absence data appeared to better approximate the three relationship categories. All coefficients of relatedness calculated from allelic frequency data were more biased for full and half siblings than estimators calculated from presence/absence data, as indicated by means that diverged from the theoretical values for full and half sibling dyads. Ritland's (1996) estimator using allelic frequency data had the

lowest standard deviations, but the differences among the means of the three relationship categories were minimal making it less useful for kin group reconstruction. Hardy's (2003) relatedness estimator had the greatest difference in distribution of values between full sibling and unrelated dyads, and was selected for partitioning full sibling clusters.

An evaluation of the type I and II error rates compared the accuracy of the seven relatedness estimators for the experimental families by using the midpoint classification cut-off between the means of the distributions of relatedness values for the two potential relationship classes (Table 3). Seven to twelve percent of known full sibling dyads were misclassified as unrelated dyads when estimators based on allelic phenotype data were examined, while a greater proportion of these dyads were misclassified with the allelic frequency calculated estimators (12–13%). Half siblings and unrelated individuals showed the highest rate of misclassification with all estimators regardless of which data format was used. The misclassification rate of unrelated individuals as full siblings was observed to be lowest with data calculated with presence/absence data (average = 8.7%) compared to the estimators calculated with allelic frequency data (average = 12.7%), and this further supported using an allelic phenotype dataset to identify full sibling clusters.

While all relatedness coefficients appeared to misclassify some proportion of full siblings, Hardy's (2003) kinship and relatedness estimators and Lynch & Milligan's (1995) estimator provided results closest to theoretical values. Hardy's (2003) kinship estimator classified 93% of full siblings correctly and 92% of unrelated individuals correctly using a full sibling  $r_{\text{partition}}$  of 0.463. To correctly classify experimentally-known full siblings 95% of the time with this kinship coefficient, a  $r_{\text{partition}}$  of 0.471 was necessary. However, this increased type II error of misclassifying unrelated individuals as full

**Table 3** Estimated misclassification rates (Type I and II error rates) calculated using the midpoint between means of pairwise value distributions for seven estimators of relatedness

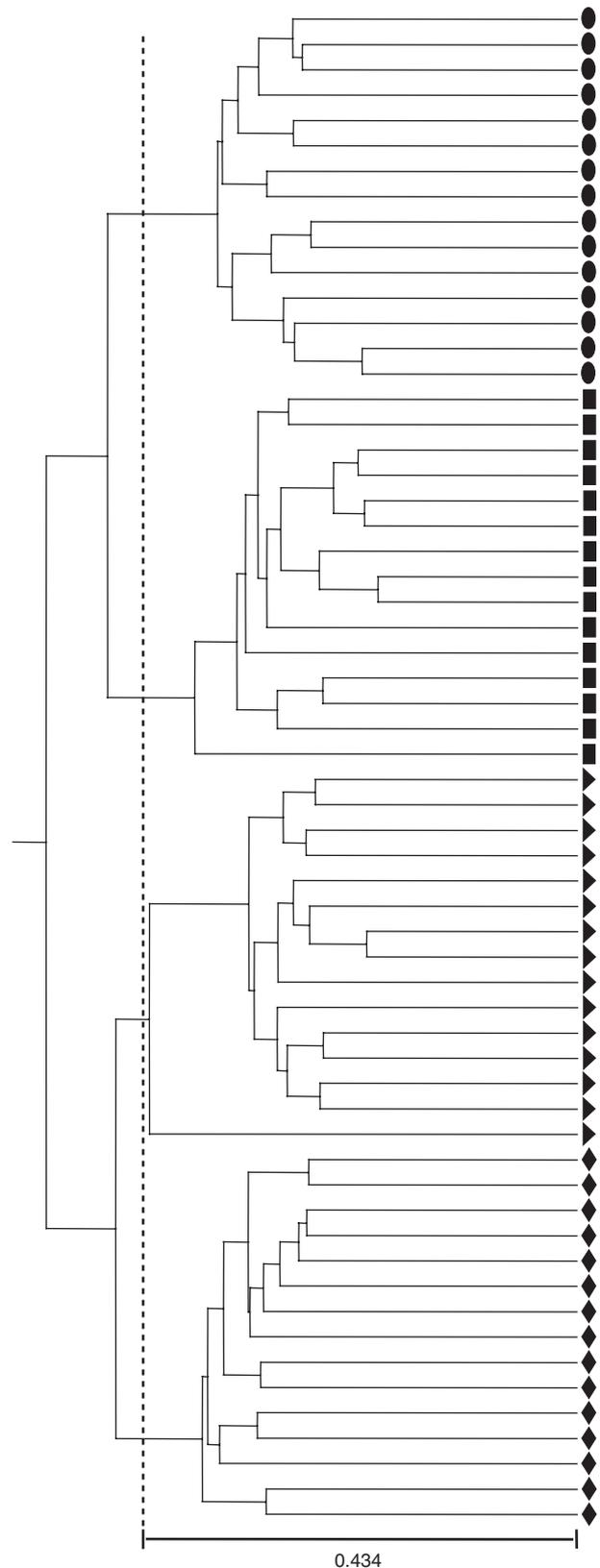
True sib relationship	Misclassified as	Dominant (Allelic phenotype)			Codominant (Allelic frequency)			
		Kinship (Hardy 2003)	Relatedness (Hardy 2003)	Relatedness (Lynch and Milligan 1994)	Ritland (1996)	Loiselle <i>et al.</i> (1995)	Streiff <i>et al.</i> (1998)	Hardy and Vekemans (1999)
Unrelated	Full sibs	0.08	0.07	0.12	0.02	0.10	0.24	0.10
Unrelated	Half sibs	0.28	0.28	0.34	0.32	0.37	0.35	0.37
Half sibs	Unrelated	0.21	0.22	0.21	0.31	0.37	0.27	0.10
Half sibs	Full sibs	0.06	0.06	0.35	0.03	0.10	0.30	0.36
Full sibs	Half sibs	0.18	0.18	0.30	0.19	0.20	0.35	0.20
Full sibs	Unrelated	0.07	0.07	0.12	0.12	0.13	0.13	0.13

siblings to 11%. Hardy's (2003) relatedness estimated classified 93% of full siblings correctly and 93% of unrelated individuals correctly using a  $r_{\text{partition}}$  of 0.434. A  $r_{\text{partition}}$  of 0.451 was necessary to correctly classify 95% of experimental full siblings with this relatedness coefficient, although this increased the type II error of misclassified experimental unrelated individuals as full siblings to 10%. Lynch & Milligan's (1995) relatedness estimated classified 88% of full siblings correctly and 88% of unrelated individuals correctly using a  $r_{\text{partition}}$  of 0.357. A  $r_{\text{partition}}$  of 0.425 was necessary to correctly classify 95% of experimental full siblings with this relatedness coefficient, although this increased type II error of misclassifying unrelated individuals as full siblings to 37%.

#### *Experimental family reconstruction and unknown fry partitioning*

The use of the UPGMA dendrograms with the relatedness coefficient of Hardy (2003) permitted for accurate reconstruction of 15 randomly selected individuals from each of the four experimental families (Fig. 2) and unknown samples (Fig. 3) when  $r_{\text{partition}}$  equalled 0.434. For the 2002 through 2006 wild green sturgeon fry datasets between six and 17 full sibling clusters were observed. The number of kin groups was evaluated with the full sibling  $r_{\text{partition}}$  of 0.434, which was the midpoint of the full sibling and unrelated pairwise value distributions and a more conservative type I error rate ( $\alpha = 0.05$ ) full sibling  $r_{\text{partition}}$  of 0.451. Except for the year 2002 ( $P^* = 0.94$ ), the probability of encountering a kin group more than once was 1.0, and all kin clusters were observed to contain more than a single offspring (Table 4). Using the 0.434  $r_{\text{partition}}$  value, 56 full sibling clusters were discovered compared to 45 clusters being partitioned over the 5 years using the 0.451 full sibling  $r_{\text{partition}}$  value. Between one (2002) and 12 (2006) individuals comprised independent full sibling clusters using the 0.434  $r_{\text{partition}}$  value.

The nonlinear regression functions evaluated were used to estimate the maximum number of full sibling kin groups being produced in the upper 96 rkms of potential green sturgeon spawning habitat in the Sacramento River (Table 4). Over the five sampled years, the



**Fig. 2** UPGMA dendrogram of 15 randomly selected full siblings from four green sturgeon progeny arrays, where four symbols represent individuals from the four experimentally-crossed families. Individuals from the 2003 half-sibling families are represented by the triangle and diamonds. The relatedness partition equals 0.434 and is half the midpoint between distributions of full sibling and unrelated individuals from controlled crosses.

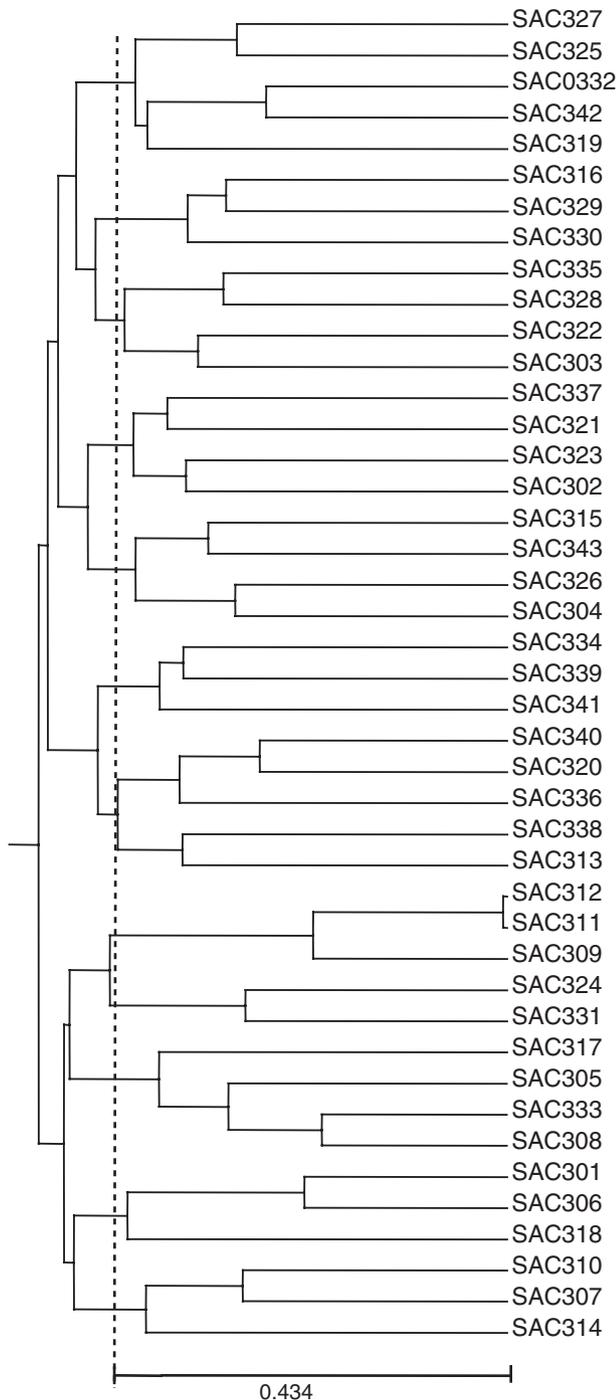


Fig. 3 UPGMA dendrogram of pairwise Hardy (2003) relatedness coefficients of 43 juvenile green sturgeon collected in 2003 at Red Bluff Diversion Dam, CA. The relatedness partition value equals 0.434 and identified 12 independent full sibling clusters.

regression analysis using Kohn's equation suggested between 6.4 to 14.1 full sibling clusters with the more conservative  $r_{\text{partition}}$ , in contrast to the 8.4 to 21.6 full sibling clusters determined with the  $r_{\text{partition}}$  of 0.434.

The three rarefaction equations estimated different asymptote (Fig. 4). Using Eggert *et al.*'s (2003) equation, the regression asymptote suggested a range of 6.4 to 16.3 full sibling clusters. Using the more conservative full sibling cluster  $r_{\text{partition}}$  of 0.451, the observed range of full sibling clusters narrowed to a minimum of 5.3 and maximum of 11.3. The regression asymptote of Chessel's (Valiere 2002) ranged from 6.2 to 13.5 families, and decreased to ranging from 5.5 to 10.3 with the 0.434  $r_{\text{partition}}$  value. We further examined the 95% confidence intervals of the accumulation curves to evaluate the accuracy of both  $r_{\text{partition}}$  values with the three accumulation functions. The number of families was estimated to be between 8.4 and 21.6 with Kohn's equation and the smaller  $r_{\text{partition}}$  value, and the estimates lowered with the conservative  $r_{\text{partition}}$  value to between 6.4 and 14.1 kin groups. Eggert *et al.*'s (2003) equation provided lower abundance results with both  $r_{\text{partition}}$  values. The estimated number of full sibling kin groups ranged from 6.4 to 16.0 with the lower type I error rate  $r_{\text{partition}}$  value, and ranged between 5.3 and 11.3 using the higher type I error rate  $r_{\text{partition}}$  value. Chessel's equation estimated a range of 6.2 to 13.5 families with the 0.434  $r_{\text{partition}}$  value and 5.5 to 10.3 families with the 0.451 full sibling  $r_{\text{partition}}$  value. Figure 5 displays a positive relationship between flows and estimated number of kin groups above the Red Bluff Diversion Dam.

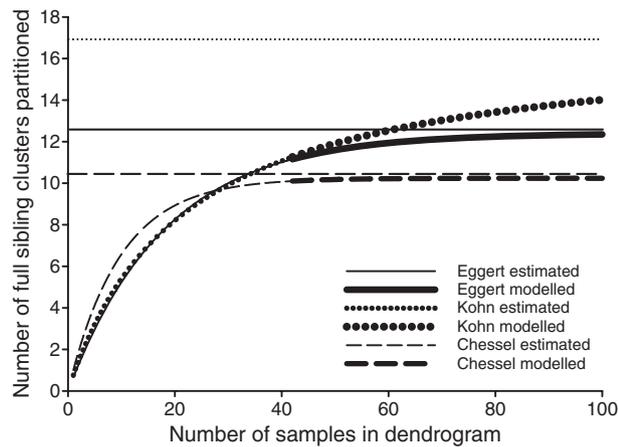
## Discussion

### *Partitioning kin groups using relatedness data*

Recent efforts to derive multilocus estimates of pairwise relatedness utilizing dominant markers suggest that approaches with these types of molecular markers can be useful in population genetic studies of organisms in which codominant microsatellite DNA markers provide a challenge for differentiating individuals from one another (Hardy 2003; Wang 2004b; Ritland 2005). The polysomic dataset described in this study provided an opportunity to evaluate how genetic data in allelic phenotype and allelic frequency formats perform and may expand the application of indirect genetic methods for estimating pairwise relatedness in complex polysomic species. The results presented in the study demonstrate the ability for presence/absence data to be of similar or increased value for describing the relationship between individual green sturgeon compared to allelic frequency data. Although other studies have used Blouin *et al.*'s (1996) method for classifying individuals into kin groups, this approach has not been used for adjusting a type I misclassification error rate for estimating more accurate results. In many capture-recapture studies, a conservative threshold of 0.05 was used to reduce the

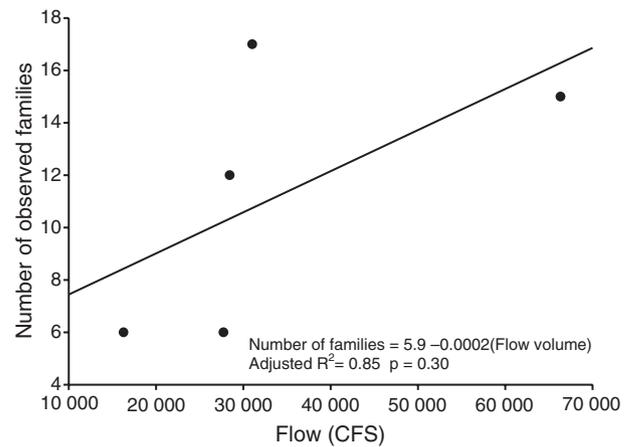
**Table 4** The observed number of full sibling kin groups and estimated maximum number of kin groups (regression asymptote) from annual collections of juvenile green sturgeon using three rarefactions equations. Both the midpoint of the means of the pairwise value distributions (0.434) and a more conservative Type 1 error rate (0.451) was used to calculate each collection's number of full sibling kin groups and estimated maximum number of kin groups. Standard error of regression is given in parenthesis

Cutoff value	p*	Observed number of kingroups	Kohn's equation		Eggert's Equation		Chessel's Equation		
			Regression asymptote	95% Confidence interval	Regression asymptote	95% Confidence interval	Regression asymptote	95% Confidence interval	
2002									
0.434	0.94	6	9.3 (1.1)	7.1, 11.5	7.0 (.6)	5.8, 8.2	7.5 (.5)	6.5, 8.5	
0.451	1	6	6.4 (.8)	4.8, 8.0	5.3 (.4)	4.5, 6.1	5.9 (.5)	5.4, 6.4	
2003									
0.434	1	12	17.0 (1.3)	14.5, 19.5	12.4 (.7)	11.0, 13.8	10.2 (.3)	9.6, 10.8	
0.451	1	11	12.1 (.7)	10.7,13.5	9.3 (.4)	8.5, 10.1	8.2 (.2)	7.8, 8.6	
2004									
0.434	1	6	8.4 (.5)	7.4, 9.4	6.4 (.3)	5.8, 7.0	6.2 (.2)	5.8, 6.6	
0.451	1	6	7.1 (.5)	6.1,8.1	5.6 (.3)	5.0, 6.2	5.5 (.2)	5.1, 5.9	
2005									
0.434	1	17	21.6 (.9)	19.8,23.4	16.0 (.5)	15.0, 17.0	13.5 (.3)	12.9, 14.1	
0.451	1	11	14.1 (.4)	13.3, 14.9	11.3 (.2)	10.9, 11.7	10.3 (.2)	9.9, 10.7	
2006									
0.434	1	15	17.9 (.6)	16.7,19.1	13.5 (.3)	12.9, 14.1	11.7 (.2)	11.3, 12.1	
0.451	1	11	13.8 (.6)	12.6,15.0	10.4 (.3)	9.8,11.0	8.5 (.2)	8.1, 8.9	



**Fig. 4** Nonlinear regression graph for three accumulation functions for number of full sibling clusters observed in UP-GMA dendrograms of juvenile green sturgeon collected in 2003. Asymptotes for each equation are horizontal lines of similar format as the individual equation's plotted line.

potential for identifying incorrect genotypes as resampled individuals (Woods *et al.* 1999; Waits *et al.* 2001; Paetkau 2003). As the likelihood of individuals sharing a genotype increases, the estimated population size is increasingly underestimated. The results using the 0.451  $r_{partition}$  value ( $\alpha = 0.05$ ) limited the potential underestimation resulting from true unrelated individuals being identified as partitioned full sibling clusters, although it



**Fig. 5** Linear regression of observed number of families partitioned annually on flow vol. past Bend Bridge (cubic feet per second (CFS)).

increased the potential for true full siblings to be mispartitioned as unrelated individuals.

The experimental family data demonstrated that partitioning half siblings from unrelated or full sibling individuals remains a challenge, although the use of polysomic microsatellite data as allele phenotypes provides sufficient accuracy for detecting full sibling vs. unrelated dyads. Ritland's (1996) method appears to provide a similar amount of error in partitioning individuals compared to Hardy's  $F_{ij}$  based upon experimental family

relatedness data (Table 2); however, the greater type I error associated with this codominant estimator made it less desirable for our analysis. Hardy's (2003) kinship and relatedness coefficients both required and account for information about inbreeding among individuals in the dataset, and this information may have increased accuracy with the estimators. While the coefficients calculated from the experimental families that were based on allele phenotypes may be more accurate than other estimators because of the power of a large number of alleles, significant problems may arise due to the bias of a limited number of alleles present in the collections of potential adults.

Numerous assumptions are necessary about mutations, null alleles, incomplete sampling of the pool of potential parents and relatedness among the parents within and among different river systems to determine the efficacy of this partitioning approach based on kinship. In fact, the allelic phenotype format of this dataset may have performed better due to lack of null alleles influencing or microsatellite genotype scoring errors associated with gene dosage interpretation of the allelic frequency data. However, the estimates of relatedness in this study are based on empirical pedigree information, and the kinship coefficient (Hardy 2003) used in this study,  $F_{ij}$ , incorporates information about the relatedness of the potential parents themselves. These pieces of information are necessary if this approach to estimating breeding population size is to be utilized among demographically independent populations. The difficulty in estimating relatedness in polyploid species often is due to a lack of this type of data, yet these were incorporated into the estimates determined in this study. While this study's comparison of multiple estimators of relatedness does provide some information to assess each algorithm's utility for kin group partitioning, the remaining assumptions require additional laboratory and simulation study. A single researcher experienced with the complexities of the microsatellite markers completed all the genotyping, although human factors have been noted to be non-negligible error generators (Bonin *et al.* 2004). Additional simulations would be useful to determine if the limited number of dyad included in calculating the distributions fully capture the range of these values in the natal populations. To approach this, it may be useful to evaluate the allele frequencies of the potential parents of Sacramento River green sturgeon, and simulate offspring based on this information. Once simulated offspring genotypes are created, distributions of relatedness values can be described for dyads of different kin relationships, which may more accurately reflect the potential values observed in each of the natal populations.

The three accumulation functions used in our analysis performed differently in identifying the maximum number of kin groups as their asymptote (Table 4). The Eggert and Chessel equations consistently attained their asymptote within the limit of the estimated kin group data. The equation suggested by Kohn *et al.* (1999) did not reach the asymptote within the estimated data in any year. In four out of five years, the range of kin groups derived from the Kohn's equation analysis was larger than with the other methods, and did not include the estimated number of kin groups. In a simulation study (Frantz *et al.* 2006), the Kohn equation overestimated population size. The analysis with Eggert's equation in four out of five years resulted in a range of kin groups that included the estimated number of kin groups above the collection point. This suggested that this function approximates the estimates of full-sibling kin groups well, and Frantz & Roper (2006) observed this equation to consistently provide accurate estimates. In contrast, the Chessel equation consistently underestimated the maximum number of kin groups in all of the collections to be smaller than the empirical number of kin groups detected in the sample. Frantz & Roper (2006) observed the Chessel equation overestimated estimates when sample sizes were small, but was accurate when sample sizes were larger.

#### *Estimated green sturgeon kin group abundance*

The partitioning of full sibling clusters containing annual fry collections using UPGMA revealed the estimated number of individuals contributing to spawning in the upper extent of the Sacramento River has remained approximately the same over the 5-year period of the study and conservatively ranged from five to 14 families. The number of partitioned kin groups above Red Bluff Diversion Dam annually increased with greater flows in the spawning reaches above Red Bluff Diversion Dam. A similar phenomenon has been observed in the sympatric white sturgeon, in which recruitment appears positively associated with higher freshwater flows through the estuary (Kohlhorst *et al.* 1991).

There may be potential reproductive consequences associated with the low numbers of green sturgeon kin groups identified annually in the upper anadromous reaches of the Sacramento River. In particular, it is possible that reproductive green sturgeon may not be able to find mates in the Sacramento River, which is hundreds of kilometres long. There are no known behavioural or physiological mechanisms that sturgeon use to locate each other, although it is possible these fish may use bioacoustic mechanisms for this purpose (Johnston & Phillips 2003). If reproductive green sturgeon are as

uncommon as these data suggest these fish may only be mating in monogamous pairings due to an inability to locate other mates. If this was the mating scheme for wild green sturgeon, then between 10 and 28 spawners would minimally contribute to kin groups above Red Bluff Diversion Dam annually. If polyandry does play a significant role in the mating behaviour of green sturgeon when they broadcast spawn, then the number of individuals contributing to spawning above Red Bluff Diversion Dam would be even lower than estimated with these approaches. This would occur because each kin group would arise from an independent monogamous pair, rather than multiple kin groups being derived from the same matriline. The possibility of this mating behaviour being used is high, and a skewed sex ratio has been observed in the Klamath River green sturgeon population (Northern DPS), where the sex ratio of adult green sturgeon was 1.0 female: 1.4 males, which was significantly different from the expected 1:1 ratio (Van Eenennaam *et al.* 2006).

Moyle (2002) suggested that between 140 and 1600 green sturgeon adults might occupy the San Francisco Bay-Delta estuary annually. While this estimate required significant assumptions regarding adult size and the ratio of white to green sturgeon captured during California Department of Fish and Game monitoring efforts and white sturgeon population estimates, the discrepancy between genetic and field estimates for the annual abundance of adults suggests either a large number of green sturgeon adults do not enter the Sacramento River or they do not migrate into the upper reaches of their spawning grounds. On the Sacramento River, spawning is believed not to occur below Hamilton City, CA (rkm 330), although the estimates of annual breeding abundance presented here do not reflect all potential spawners.

While these data conclusively demonstrate multiple families of green sturgeon in the wild, estimates of breeding population size do not have a significant positive relationship with increased flows. The lack of significance of this positive relationship may be due to the limited number of samples, which makes statistical tests unreliable (Townend 2005). It is also possible that flow does not have a significant influence on breeding population size in this portion of the Sacramento River, and escapement by reproductive adults is a result of productive ocean cycles or other physical drivers. Red Bluff Diversion Dam, a seasonal irrigation diversion, was closed during portions of the spawning season and has been noted to impact the number of adults ascending the portion of the Sacramento River (Heublein *et al.* 2009). Other mechanisms may drive the recruitment of juvenile green sturgeon in the Sacramento River including aspects of their reproductive ecology, predation, or

juvenile rearing habitat. If the sampling efforts that contributed to this study randomly sampled all families spawned above the Red Bluff Diversion Dam equivalently (possible in 2005 and 2006), it appears that a small number of large families and a larger number of small families were surveyed in the sampling. Fecundity and egg diameter have a positive relationship with the size of female green sturgeon, and Van Eenennaam *et al.* (2006) suggested older females, which are assumed to spawn multiple times, reach their peak egg production when they are older than 32 years of age. While there are no year class strength data for Southern DPS green sturgeon, these data suggest not all spawners contribute equally to the population. It is possible that many younger females are contributing smaller clutches to the annual productivity of green sturgeon on the Sacramento River, while only a small number of the spawners are from older age classes. This possibility is consistent with empirical observations of anadromous Gulf sturgeon, where a majority of spawning is dominated by younger individuals and large, older fish are rare (Sulak & Randall 2002).

The probability of recapturing a kin group was high in every year. By evaluating the location of the modelled asymptote, it is possible to determine how many samples may be necessary for describing the maximum number of kin groups in a collection. In 2003, the Eggert and Chessel equations reached an asymptote within 60 samples, although we did not have the number of samples to assess if results followed the accumulation curve. These relationships are similar to what were observed in the other years, thus it is possible that this method, which takes advantage of partitioning kin groups, can effectively characterize the reproductive population of rare species with minimal sized collections of tissue samples. The collections with the best sampling regimes were in 2003, 2005, and 2006, and when samples were added to the annual dendrograms in the order they were collected, all kin groups were sampled within the first pulse of greens sturgeon emigrating past Red Bluff Diversion Dam.

The approach used in this study could form the basis for an effective method for estimating the abundance of breeding green sturgeon, and additional simulations investigating underlying assumptions are important to validating this framework. The use of experimental families, such as those employed in this study, is necessary for converting relatedness values among wild individuals when their potential parents have not been sampled. Partitioning juveniles by reconstructing their pedigree is a valuable approach, since it is often not possible to sample adults that mate cryptically, broadcast spawn, or are rare. The Blouin *et al.* (1996) method of identifying classification error provided an analogous

framework for partitioning individual genotypes into categories of the same relatedness as probability of identity does for traditional genetic 'capture-recapture' studies. This approach has applications when mating individuals are unknown, yet it is possible to survey a portion of the fry or juvenile production as a means of measuring the annual spawner abundance. In order to assess the importance of the upper portions of spawning habitat to recruitment of juveniles in the green sturgeon Southern DPS, additional field studies should evaluate green sturgeon spawning in other portions of the Sacramento River through collections of juveniles in these areas. This framework also provides a potential mechanism for studying recruitment dynamics of species by combining individual age data with genotype data. In this case, it may be possible to reconstruct kin groups within cohorts of older juveniles and adults to assess the contribution of distinct families and the number of reproductive adults contributing to older cohorts, which have influenced the demography of the population.

### Acknowledgements

We thank D. Hillemeier and associates of the Yurok Tribal Fisheries Program, S. Doroshov and J. van Eenennaam of UC Davis, B. Poytress and associates of the Red Bluff U.S. Fish and Wildlife Service Office, and M. Gingras and associated of the California Department of Fish and Game for assistance with sampling. Earlier versions of this paper benefited from constructive comment from P. Moyle, W. Ardren, E. Anderson, M. Hansen and two anonymous reviewers. Discussions with O. Hardy aided in calculation of kinship coefficients. This study has been supported by funds from CALFED (#ERP-02D-P57, 98-C15, 99-F105), Anadromous Fish Restoration Program (11332-1-G005), and NOAA-National Marine Fisheries Service (AB133F075E3873). JAI received financial assistance from UC Davis Ecology Graduate Group Block Grants, Hart, Cole, Goss Summer Research Fellowships, and Butler and Walker University Fellowships to complete this study.

### References

- Almudevar A, Field C (1999) Estimation of Single-Generation Sibling Relationships Based on DNA Markers. *Journal of Agricultural Biological and Environmental Statistics*, **4**, 136–165.
- Becher SA, Steinmetz K, Weising K, *et al.* (2000) Microsatellites for cultivar identification in *Pelargonium*. *Theoretical and Applied Genetics*, **101**, 643–651.
- Beyer J, May B (2003) A graph-theoretic approach to the partition of individuals into full-sib families. *Molecular Ecology*, **12**, 2243–2250.
- Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *TRENDS in Ecology and Evolution*, **18**, 503–511.
- Blouin MS, Parsons M, Lacaille V, *et al.* (1996) Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology*, **5**, 393–401.
- Bonin A, Bellemain E, Bronken Eidesen P, *et al.* (2004) How to track and assess genotyping error in population genetic studies. *Molecular Ecology*, **13**, 3261–3273.
- Caujape-Castells J, Baccarani-Rosas M (2005) *Transformer-3*: a program for the analysis of molecular population genetic data. EXEGEN software and Jardín Botánico Canario 'Viera y Clavijo'.
- Eggert LS, Eggert JA, Woodruff DS (2003) Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. *Molecular Ecology*, **12**, 1389–1402.
- Felsenstein J (2005) PHYLIP, Phylogeny Inference Package (version 3.6.6). Available from <http://evolution.genetics.washington.edu/phylip.html> (Accessed 20 September 2006).
- Frantz AC, Fack F, Muller CP, *et al.* (2006) Faecal DNA typing as a tool for investigating territorial behaviour of badgers (*Meles meles*). *European Journal of Wildlife Research*, **52**, 138–141.
- Frantz AC, Roper TJ (2006) Simulations to assess the performance of different rarefaction methods in estimating population size using small datasets. *Conservation Genetics*, **7**, 315–318.
- Goodnight KF, Quller DC (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology*, **8**, 1231–1234.
- Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.
- Gottelli D, Wang JL, Bashir S, *et al.* (2007) Genetic analysis reveals promiscuity among female cheetahs. *Proceedings of the Royal Society B-Biological Sciences*, **274**, 1993–2001.
- Guillot G, Leblois R, Coulon A, *et al.* (2009) Statistical methods in spatial genetics. *Molecular Ecology*, **18**, 4734–4756.
- Hardy OJ (2003) Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology*, **12**, 1577–1588.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology*, **2**, 618–620.
- Hardy OJ, Vekeman X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83**, 145–154.
- Hedrick PW (2000) *Genetics of Populations*, 2nd edn. Jones and Bartlett Publishers, Sudbury.
- Herbinger CM, O'Reilly PT, Verspoor E (2006) Unraveling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. *Molecular Ecology*, **15**, 2261–2275.
- Heublein JC, Kelly JT, Crocker CE, *et al.* (2009) Migration of green sturgeon, *Acipenser medirostris*, in the Sacramento River. *Environmental Biology of Fishes*, **84**, 245–258.
- Israel JA, Bando KJ, Anderson EC, *et al.* (2009) Polyploid microsatellite data reveal stock complexity among estuarine North American green sturgeon (*Acipenser medirostris*). *Canadian Journal of Fisheries and Aquatic Sciences*, **66**, 1491–1504.
- Israel JA, Cordes JF, Blumberg MA, *et al.* (2004) Geographic patterns of genetic differentiation among collections of green sturgeon. *North American Journal of Fisheries Management*, **24**, 922–931.

- Johnston CE, Phillips CT (2003) Sound production in sturgeon *Scaphirhynchus albus* and *S. platyrhynchus* (Acipenseridae). *Environmental Biology of Fishes*, **68**, 59–64.
- Kohlhorst DW, Botsford LW, Brennan JS, *et al.* (1991) Aspects of the structure and dynamics of an exploited Central California population of white sturgeon (*Acipenser transmontanus*). In: *Acipenser* (ed. Williot P), pp. 277–293. CEMAGREF, Bordeaux.
- Kohn MH, York EC, Kamradt DA, *et al.* (1999) Estimating population size by genotyping faeces. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **266**, 657–663.
- Konovalov DA, Manning C, Henshaw MT (2004) KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Molecular Ecology Notes*, **4**, 779–782.
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Loiselle BA, Sork VL, Nason J, *et al.* (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.
- Ludwig A, Belfiore NM, Pitra C, *et al.* (2001) Genome duplication events and functional reduction of ploidy levels in Sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). *Genetics*, **158**, 1203–1215.
- Lukacs PM, Burnham KP (2005) Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, **14**, 3909–3919.
- Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.
- Makinen T, Panova M, Andre C (2007) High levels of multiple paternity in Littorina saxatilis: hedging the bets? *Journal of Heredity*, **98**, 705–711.
- Moyle P (2002) *Inland Fishes of California*, 2nd edn. University of California Press, Berkeley.
- Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology*, **12**, 1375–1387.
- Pearse DE, Eckerman CM, Janzen FJ, *et al.* (2001) A genetic analogue of 'mark-recapture' methods for estimating population size: an approach based on molecular parentage assessments. *Molecular Ecology*, **10**, 2711–2718.
- Ritland K (2005) Multilocus estimation of pairwise relatedness with dominant markers. *Molecular Ecology*, **14**, 3157–3165.
- Ritland K (1996) A marker-based method for pairwise relatedness and individual inbreeding coefficients. *Genetical Research*, **67**, 175–185.
- Rodzen JA, Famula TR, May B (2004) Estimation of parentage and relatedness in the polyploid white sturgeon (*Acipenser transmontanus*) using a dominant marker approach for duplicated microsatellite loci. *Aquaculture*, **232**, 165–182.
- Smith BR, Herbinger CM, Merry HR (2001) Accurate partition of individuals into full-sib families from genetic data without parental information. *Genetics*, **158**, 1329–1338.
- Streiff RT, Labbe R, Bacilieri H, *et al.* (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology*, **7**, 317–328.
- Sulak K, Randall M (2002) Understanding sturgeon life history: Enigmas, myths, and insights from scientific studies. *Journal of Applied Ichthyology*, **18**, 519–528.
- Thrall PH, Young A (2000) AUTOTET: A program for analysis of autotetraploid genotypic data. *Journal of Heredity*, **91**, 348–349.
- Townend J (2005) *Practical Statistics for Environmental and Biological Scientists*. John Wiley & Sons, Ltd, West Sussex.
- Valiere N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377–379.
- Van Eenennaam JP, Linares J, Doroshov SI, *et al.* (2006) Reproductive conditions of the Klamath River green sturgeon. *Transactions of the American Fisheries Society*, **135**, 151–163.
- Van Eenennaam JP, Linares-Casenave J, Deng X, *et al.* (2005) Effect of incubation temperature on green sturgeon embryos, *Acipenser medirostris*. *Environmental Biology of Fishes*, **72**, 145–154.
- Van Eenennaam JP, Webb MAH, Deng X, *et al.* (2001) Artificial spawning and larval rearing of Klamath River green sturgeon. *Transactions of the American Fisheries Society*, **130**, 159–165.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.
- Wang JL (2004a) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963–1979.
- Wang J (2004b) Estimating pairwise relatedness from dominant genetic markers. *Molecular Ecology*, **13**, 3169–3178.
- Waits L, Luikart G, Taberlet P (2001) Noninvasive genetic samples: look before you leap. *Trends in Ecological Evolution*, **14**, 323–327.
- Woods JG, Paetkau D, Lewis D, *et al.* (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, **27**, 616–627.

---

This study was part of J.A.I.'s PhD dissertation on conservation genetics of green sturgeon. His interests focus on the integration of laboratory studies and population monitoring for informing species recovery implementation, aquatic conservation planning and adaptive ecosystem management. B.M. is an Adjunct Professor of Animal Science and Director of the Genomic Variation Laboratory (GVL). Researchers at the GVL use molecular genetic techniques and analyses to address a diverse array of questions related to the conservation of threatened and endangered species and the genetic improvement of aquacultural taxa.

---