Polyploid microsatellite data reveal stock complexity among estuarine North American green sturgeon (*Acipenser medirostris*)

Joshua A. Israel, K. Jun Bando, Eric C. Anderson, and Bernie May

Abstract: North American green sturgeon (*Acipenser medirostris*) display the distinctive behavior of long ocean migrations along the west coast punctuated by summer residence in estuaries; however, little is known about stock composition in these tidal environments. Pairwise comparisons and genetic clustering analysis were used to characterize the two green sturgeon Distinct Population Segments (DPSs) from 20 collections using eight tetrasomic and two disomic microsatellite loci. The observed pattern of green sturgeon DPS composition among five estuaries in California, Oregon, and Washington was supported with assignment testing approaches utilizing the same genotypic data in codominant polysomic and pseudodominant allele phenotype formats. The majority of individuals in northern DPS estuaries originated from the threatened Southern DPS, except in Winchester Bay and Grays Harbor. We detected few Northern DPS green sturgeon in San Pablo Bay, the principal Southern DPS estuary, supporting that green sturgeon preferentially disperse north once they enter their coastal migration. Our genetic findings suggest that stock complexity in green sturgeon is pervasive and support precautionary, interjurisdictional approaches for managing green sturgeon beyond rigid, regulatory boundaries.

Résumé : L'esturgeon vert (*Acipenser medirostris*) d'Amérique du Nord possède un comportement particulier qui comprend de longues migrations en mer le long de la côte ouest interrompues par des séjours d'été dans les estuaires; on connaît cependant peu la composition des stocks dans ces milieux sous influence de la marée. Des comparaisons appariées et une analyse de groupement génétique nous ont servi à caractériser deux segments distincts de population (DPS) d'esturgeons verts à partir de 20 échantillons et à l'étude de huit locus microsatellites tétrasomiques et deux disomiques. Le patron obtenu de composition des DPS dans cinq estuaires de Californie, d'Oregon et du Washington est confirmé par des méthodes d'essais d'attribution d'après les mêmes données génotypiques dans des formats de phénotypes d'allèles polysomiques codominants et pseudodominants. La plupart des individus des estuaires du DPS du nord proviennent du DPS menacé du sud, à l'exception de ceux de la baie Winchester et de Grays Harbor. Nous avons décelé peu d'esturgeons verts du DPS du nord dans la baie de San Pablo, l'estuaire principal du DPS du sud, ce qui appuie l'hypothèse qui veut que les esturgeons verts se dispersent de préférence vers le nord une fois qu'ils commencent leur migration côtière. Nos observations génétiques indiquent que la complexité du stock est générale et elles nous amènent à recommander des approches prudentes couvrant les diverses juridictions et dépassant les frontières rigides de la règlementation pour la gestion de l'esturgeon vert.

[Traduit par la Rédaction]

Introduction

Information about the stock composition of anadromous sturgeon aggregating in coastal bays and estuaries is very limited, yet essential for reducing risks associated with port dredging and development, coastal and estuarine fisheries, and habitat modification impacting sturgeon species in these waters. Adult North American green sturgeon (*Acipenser medirostris*) spawn in deep, cool, swift freshwater habitats and juveniles rear in natal rivers and estuaries before emigrating to the ocean in their second or third year (Allen and Cech 2007). Subadult green sturgeon make extensive oceanic migration along the Pacific coastal shelf (Erickson and Hightower 2007) northward in the fall and aggregate in bays and estuaries in the summer and fall (Moser and Lindley 2007). These estuarine aggregations do not seem related to spawning migration, since most aggregations do not occur in estuaries of spawning rivers, and green sturgeon appear to return south in the spring to enter spawning rivers (Lindley et al. 2008). Recent studies have suggested that de-

Received 12 September 2008. Accepted 30 March 2009. Published on the NRC Research Press Web site at cjfas.nrc.ca on 27 August 2009.
J20774

Paper handled by Associate Editor Michael Hansen.

J.A. Israel,[1] K.J. Bando, and B. May. Genomic Variation Laboratory, Department of Animal Science, 1 Shields Avenue, University of California, Davis, CA 95616, USA.
E.C. Anderson.[2] Southwest Fisheries Science Center, 110 Shaffer Road, Santa Cruz, CA 95060, USA.

[1]Corresponding author (e-mail: jaisrael@ucdavis.edu).
[2]Present address: Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA.

spite extensive marine migrations, anadromous sturgeon exhibit notable patterns of possibly stock-specific movement and aggregation (Waldman et al. 1996; Dugo et al. 2004). Thus, it is possible that green sturgeon migration patterns are distinct among estuaries and the relationship of these migratory patterns to the individuals' population of origin remains unknown. Recognition of stock complexity at local scale is essential for effective fisheries management, stock assessment, and recovery planning (Stephenson 1999; Cadrin et al. 2005).

North American green sturgeon range from Ensenada, Mexico, to the Bering Sea, Alaska (Moyle 2002; Colway and Stevenson 2007). Green sturgeon are currently known to spawn in two large California basins in the Sacramento and Klamath rivers and the Rogue River in Oregon. Late sexu ty (approximately 15 years of age; Van Eenennaam 2006), iteroparity, and extensive marine migrations green sturgeon to numerous threats throughout their ory. Aquatic ecosystems are strained by water manag and similar to other diadromous fish populations, g sturgeon have been impacted by flow and temperature rations in spawning habitats caused by the construc of dams, degradation of estuarine rearing and foraging tats by levee construction, dredging, and invasive speci and trawl fishing on migratory routes along the western c of North America (Adams et al. 2007; Erickson and W b 2007; Lackey 2009). The Sakhalin sturgeon, a sibling sp cies in the Asian Pacific, has been extirpated from multip spawning rivers and only persists at extremely low abundan es (Shmigirilov et al. 2007).

Table 1. Location, sampling dates, collecting organization, sample size (*n*), collection site (measured in river kilometre distance (rkm) from the mouth), and fish size range (total length) for 1196 green sturgeon samples.

Collection location	Code	Dates	Organization	*n*	Collection site	Total length range (mm)
Sacramento River, California	**SAC**	June–July 2002	USFWS	18	rkm 388.8	26–52
		June–July 2003	USFWS	43	rkm 388.8	23–30
		May–July 2004	USFWS	24	rkm 388.8	24–41
		June–July 2005	USFWS	89	rkm 388.8	25–41
		June–July 2006	USFWS	92	rkm 388.8	25–42
San Pablo Bay, California	SPB	August–September 2001	CDFG	114	Estuary	600–1830
		April–May and August–September 2004	UCD	105	Estuary	690–2030
Klamath River, California	**KLA**	April–June 1998	YTFP	60	rkm 0.0–80.0	1550–2110
		April–June 2001	YTFP	29	rkm 0.0–80.0	1450–2400
		April–May 2003	YTFP	35	rkm 0.0–80.0	1460–2230
Rogue River, Oregon	**ROG**	May–July 2000	ODFW	34	rkm 1.8–104.6	1470–2250
		April–September 2002	ODFW	45	rkm 12.0–39.2	1530–2160
		April 2003	ODFW	34	rkm 13.6–39.2	1450–2130
Winchester Bay, Oregon	WIN	August 2000	ODFW	97	rkm 14.4–19.2	760–2110
		June–July 2002	ODFW	22	rkm 14.4–20.8	660–2080
Columbia River, Washington	COL	October 1995	WDFW	32	rkm 0.0- 32.0	1190–1590
		1999	ODFW	49	Lower Columbia River	na
		July–August 2004	WDFW	94	rkm 18.8–35.2	980–2220
Willapa Bay, Washington	WIL	July–September 2003	WDFW	98	Areas WF2GW, WF2GE, and WF02J	1050–1950
Grays Harbor, Washington	GHR	June–July 2005	WDFW	82	Areas WFO2B and WFO2D	990–2400

Note: Population codes for spawning rivers are in bold type. USFWS, US Fish and Wildlife Service; CDFG, California Department of Fish and Game; UCD, University of California, Davis; YTFP, Yurok Tribal Fisheries Program; ODFW, Oregon Department of Fish and Wildlife; WDFW, Washington Department of Fish and Wildlife. na, not available.

Table 2. PCR conditions, genotyping dilutions optimized for electrophoresis, and PCR success rate for each locus.

Locus	Primer source(s)	PCR profile	Optimized conditions	PCR dilution	Ladder (bp)	PCR success rate (%)
Spl120b	McQuown et al. 2000; Israel et al. 2004	Fastart	2.0 mmol/L MgCl[a]	1 to 2	400	95.9
Aox27	King et al. 2001	Touchdown	1.25 mmol/L MgCl[b]	1 to 8	400	89.3
AfuG43	Welsh et al. 2003	Promega	1.75 mmol/L MgCl,[c] T_A = 59 °C	1 to 2	400	98.8
Spl106	McQuown et al. 2000	Touchdown	2.0 mmol/L MgCl[b]	1 to 2	400	98.0
AfuG135	Welsh et al. 2003	Promega	2.0 mmol/L MgCl,[c] T_A = 52 °C	1 to 8	400	94.5
Spl101b	Israel et al. 2004	Promega	T_A = 54 °C[c]	1 to 8	400	95.6
AfuG41	Welsh et al. 2003	Touchdown	2.0 mmol/L MgCl[b]	1 to 8	400	95.7
AfuG247	Welsh et al. 2003	Fastart	2.0 mmol/L MgCl[a]	1 to 2	400	94.7
An76	Zane et al. 2002	Fastart	T_A = 50 °C[a]	No dilution	400	90.0
As007	Zhu et al. 2005	Promega	2.5 mmol/L MgCl[c]	1 to 4	500	94.3

[a]95 °C for 4 min; 30 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min; 72 °C for 5 min.

[b]95 °C for 1 min; 15 cycles of 95 °C for 30 s, 65 °C for 1min (–1 °C per cycle), and 72 °C for 1 min; 15 cycles of 95 °C for 30 s, 50 °C for 1 min, and 72 °C for 1 min.

[c]95 °C for 1 min 30 s; 40 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min; 72 °C for 5 min.

In response to population declines, the National Marine Fisheries Service conducted a biological review of green sturgeon. One of the review's most important findings was that there are two spawning populations that appear to be reproductively isolated (Israel et al. 2004; Adams et al. 2007). The southern group, called the Southern Distinct Population Segment (DPS), consists of fish that spawn only in the Sacramento River. It was listed in 2006 as threatened under the *Endangered Species Act*. The Northern DPS, a National Marine Fisheries Service Species of Concern, consists of green sturgeon that spawn in rivers north of and including the Eel River in California. It is believed that fish in the Southern DPS are considerably less abundant than in the Northern DPS rivers, where more spawning habitat is available, harvest monitoring indicates hundreds of annual adult spawners, and there is minimal loss to entrainment (Adams et al. 2007).

Understanding how the behavior and population structure of green sturgeon influence the species' responses to impacts in diverse riverine, estuarine, and ocean habitats is challenging. Recent ultrasonic tagging studies have been successful in illuminating migration patterns of the species, although these studies alone cannot answer questions about the impacts of fishing or environmental alterations upon

Table 3. Estimates of observed (H_o) and expected heterozygosity (H_e) and number of alleles per locus (A) in each collection.

	Double reduction observed									Tetrasomic					
	AfuG41			*An76*			*AfuG135*			*AfuG43*			*Spl106*		
Collection	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A
SAC 2002	0.59	0.64	9	0.32	0.75	12	0.41	0.55	6	0.75	0.64	6	0.19	0.25	3
SAC 2003	0.61	0.70	11	0.30	0.73	12	0.39	0.49	6	0.67	0.62	7	0.16	0.24	6
SAC 2004	0.69	0.74	9	0.46	0.73	11	0.22	0.32	5	0.51	0.56	4	0.06	0.06	3
SAC 2005	0.57	0.68	9	0.33	0.69	14	0.31	0.42	9	0.57	0.61	6	0.10	0.11	4
SAC 2006	0.62	0.68	11	0.30	0.75	16	0.32	0.46	10	0.68	0.69	7	0.19	0.26	6
SPB 2001	0.59	0.71	13	0.28	0.75	16	0.38	0.51	14	0.63	0.64	7	0.19	0.25	10
SPB 2004	0.66	0.71	10	0.28	0.69	13	0.36	0.46	9	0.63	0.62	6	0.14	0.15	6
KLA 1998	0.64	0.80	13	0.51	0.80	17	0.53	0.82	13	0.76	0.74	9	0.48	0.46	6
KLA 2001	0.78	0.80	13	0.43	0.71	13	0.62	0.77	11	0.47	0.50	6	0.64	0.62	7
KLA 2003	0.63	0.79	12	0.38	0.78	13	0.55	0.82	12	0.75	0.72	7	0.37	0.36	3
ROG 2000	0.76	0.80	13	0.30	0.80	15	0.60	0.82	13	0.72	0.71	8	0.53	0.57	5
ROG 2002	0.79	0.79	12	0.41	0.82	18	0.60	0.80	12	0.69	0.76	9	0.40	0.38	4
ROG 2003	0.61	0.82	12	0.63	0.79	9	0.54	0.81	11	0.75	0.73	8	0.27	0.32	3
WIN 2002	0.76	0.77	10	0.41	0.81	13	0.61	0.80	10	0.71	0.75	7	0.45	0.49	6
WIN 2000	0.64	0.76	17	0.31	0.75	20	0.44	0.65	13	0.68	0.70	9	0.23	0.24	8
COL 1995	0.64	0.72	11	0.23	0.69	11	0.31	0.54	9	0.65	0.65	6	0.15	0.15	4
COL 1999	0.67	0.78	13	0.43	0.77	21	0.32	0.54	10	0.65	0.65	7	0.15	0.14	5
COL 2004	0.63	0.74	14	0.30	0.79	19	0.40	0.60	11	0.63	0.62	7	0.10	0.10	4
WIL 2003	0.60	0.75	14	0.46	0.80	19	0.41	0.57	10	0.68	0.68	8	0.13	0.15	4
GHR 2005	0.63	0.80	14	0.37	0.81	19	0.57	0.77	14	0.69	0.70	11	0.12	0.12	5

Southern and Northern DPS green sturgeon in specific areas (Lindley et al. 2008). Typically, green sturgeon have been tagged in estuaries, where they are easy to capture. A number of recaptures and acoustic detections of green sturgeon tagged in Northern DPS estuaries and Southern DPS rivers and these locations indicate migration by individuals between the two DPSs (Miller 1972; Lindley et al. 2008). However, information derived from the tagging effort is limited by not knowing the population of origin for a majority of these fish and the rates of marking and recapture or detection, in the case of ultrasonic tagging studies.

Genetic stock identification (GSI) studies can provide information about the origin of green sturgeon in aggregations and only require a single contact with the sample. However, green sturgeon population studies are complicated by their degenerating tetraploid genome. Most microsatellite markers used in this study carry four copies of most genes, although some seem to have decayed to amplify two copies of a gene. Tetrasomy makes it very difficult to verify that the assumptions underlying standard statistical genetic methods are satisfied; however, these methods can still yield accurate results. Salmonid genomes are also tetraploid derivatives (Wright et al. 1983), and the treatment and assumptions surrounding the use of isolocus allozyme frequency data for population differentiation and GSI studies were often different among studies (Marsden et al. 1989; Scribner et al. 1998; Seeb et al. 2000). Applications of GSI (Beacham et al. 1987; Smouse et al. 1990) and other methods of genetic assignment testing (Paetkau et al. 1995) are now commonplace with diploid microsatellite data. In diploid species, interpreting microsatellite data and applying the statistical models underlying GSI follow the relatively simple and well-understood laws of Mendelian inheritance. This simplicity is lost in polyploid organisms such as green sturgeon, and without careful multigenerational studies, the statistical model describing how tetrasomic genes are inherited cannot be known with confidence. Accordingly, the application of GSI to polysomic microsatellite data has been somewhat restricted.

The objectives of this study were to (*i*) test the use of genetic assignment testing programs suitable for tetrasomic microsatellite data and (*ii*) evaluate the stock composition of green sturgeon in estuarine aggregations. Genetic assignment testing methods were examined using two different scoring models (a dominant versus a fully-resolved codominant model). Further, we assessed the accuracy of these GSI estimates by resampling entire individuals into simulated baselines and mixtures. This procedure allowed us to challenge our methods with simulated data that were faithful to the tetrasomic inheritance mechanism of each locus, without that mechanism being exactly known.

Green sturgeon in the Southern and Northern DPSs were previously shown to be genetically differentiated at six microsatellite DNA loci (Israel et al. 2004). In this study, we used these microsatellite markers along with four others to identify Northern and Southern DPS green sturgeon individuals in nine collections of green sturgeon taken from five estuaries in California, Oregon, and Washington to evaluate two estuarine aggregation hypotheses. The null hypothesis that green sturgeon in estuaries originated primarily from their most proximal geographic population was examined. Alternatively, the possibility that estuarine green sturgeon were observed in proportion to the purported abundance of the two DPSs was also evaluated.

Materials and methods

Collection of DNA samples, polymerase chain reaction (PCR), and gel electrophoresis

Green sturgeon samples were collected from eight west coast locations in the United States (Fig. 1). Collections of freshwater samples were assumed to represent spawning

									Disomic					
Spl101b			*AfuG247*			*As007*			*Aox27*			*Spl120*		
H_o	H_e	*A*	H_o	H_e	*A*	H_o	H_e	*A*	H_o	H_e	*A*	H_o	H_e	*A*
0.68	0.72	9	0.46	0.53	4	0.71	0.76	8	0.46	0.37	4	0.68	0.67	6
0.66	0.67	12	0.59	0.67	9	0.76	0.78	8	0.41	0.54	4	0.61	0.78	8
0.75	0.77	12	0.51	0.64	6	0.65	0.77	8	0.44	0.36	4	0.67	0.70	7
0.84	0.82	15	0.59	0.68	9	0.70	0.77	13	0.67	0.51	4	0.33	0.71	5
0.72	0.80	16	0.55	0.68	10	0.75	0.83	9	0.44	0.46	5	0.64	0.69	8
0.80	0.83	22	0.48	0.61	10	0.74	0.82	12	0.46	0.48	4	0.64	0.73	8
0.79	0.80	20	0.57	0.64	8	0.76	0.80	10	0.60	0.68	4	0.84	0.83	9
0.83	0.83	13	0.49	0.68	13	0.69	0.77	12	0.62	0.65	5	0.86	0.84	13
0.78	0.82	9	0.50	0.60	8	0.72	0.74	8	0.57	0.62	4	0.61	0.83	8
0.79	0.78	10	0.50	0.64	8	0.77	0.80	11	0.74	0.65	3	0.76	0.84	11
0.82	0.82	12	0.48	0.65	11	0.65	0.83	13	0.71	0.63	4	0.76	0.81	12
0.84	0.82	10	0.60	0.66	9	0.79	0.83	10	0.73	0.67	4	0.47	0.79	11
0.73	0.76	10	0.43	0.70	9	0.79	0.79	9	0.58	0.66	4	0.82	0.85	11
0.88	0.84	10	0.55	0.71	8	0.74	0.75	8	0.53	0.55	4	0.50	0.79	12
0.75	0.81	17	0.53	0.65	12	0.77	0.82	11	0.55	0.54	6	0.90	0.83	11
0.87	0.85	16	0.59	0.64	8	0.70	0.82	8	0.45	0.39	3	0.79	0.79	9
0.65	0.70	17	0.46	0.70	6	0.74	0.77	9	0.42	0.47	3	0.68	0.80	11
0.77	0.78	19	0.55	0.60	9	0.72	0.77	10	0.56	0.52	3	0.76	0.73	8
0.82	0.83	20	0.58	0.61	10	0.72	0.79	12	0.47	0.44	5	0.73	0.78	10
0.73	0.84	15	0.54	0.71	13	0.74	0.80	12	0.49	0.52	4	0.73	0.79	9

populations, while estuarine collections were of unknown origin. Spawning populations were represented by larvae and fry (age 0+) samples from the Sacramento River and adult samples (unknown ages) from the Klamath and Rogue rivers (Table 1). Green sturgeon occupying fresh and estuarine waters before the onset of maturation are referred to as juveniles, while fish between the onset of maturation and full maturity that have entered coastal and estuarine waters are called subadults (Klimley et al. 2007). Adult sturgeon are reproductively mature fish and can be found in all aquatic environments. Fin clips of individuals were either dried or preserved in 95% ethanol and stored at room temperature. DNA was isolated using the Wizard SV DNA extraction kit (Promega). Because green sturgeon may be misidentified as white sturgeon, the taxonomic identity of all samples was verified using a green sturgeon specific restriction fragment length polymorphism (Israel et al. 2004).

In addition to the six microsatellite DNA loci previously described in Israel et al. (2004), we used four additional loci: *AfuG41* and *AfuG247* (Welsh et al. 2003), *An76* (Zane et al. 2002), and *As007* (Zhu et al. 2005). These 10 loci (Table 2) were amplified by PCR in each of the samples and genotyping dilutions optimized for electrophoresis on the BaseStation DNA fragment analyzer (MJ Research/Bio-Rad, Inc.). PCR success rates for each locus are reported in Table 2. A 0.5 μmol/L solution containing the forward primer end-labeled with one of three PRISM fluorophores, NED, VIC, or 6-FAM (Applied Biosystems Inc.), was included in each primer set. One microlitre of diluted PCR product for each locus was pooled with products from two other labeled primers and combined with 2 μL of water, 1.45 μL of 100% deionized formamide (Sigma), 0.5 μL of blue dextran loading dye, and 0.05 μL of ABI GeneScan-400 Rox internal size standard. Samples were denatured at 95 °C for 3 min and chilled on ice. One microlitre aliquots of PCR product and size standard mixtures were loaded manually into a 48-well, 75 μm, 5.5% denaturing polyacrylamide gel and size fractionated by electrophoresis in 1× TBE running buffer with a 2 min prerun at 1900 V, a 30 s constant injection at 4000 V, and a 6000 scan collection run at 2600 V. PCR products were analyzed and genotypic data were generated using Cartographer (version 1.2.6g) DNA fragment analysis software. Multiple individuals were rerun on each gel to standardize intergel allele size and and validate gene dosage scoring.

Gene dosages for the eight tetrasomic microsatellite DNA markers were interpreted by analysis of relative peak intensity from the genotyping electropherograms, similar to other sturgeon studies with polysomic markers (McQuown et al. 2000; Pyatskowit et al. 2001; Welsh et al. 2003). If three peaks were present, the largest peak was assigned two doses. If two peaks were present and equal in size, each peak was scored as representing two doses. When two peaks were present but not equal in size, the largest peak was interpreted as presumably three doses of its corresponding allele. 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. Since scoring gene dosage of tetrasomic loci may not be 100% accurate for estimating allele and genotype frequencies (De Silva et al. 2005; Luo et al. 2005), we also conducted analyses based merely on the presence or absence of different alleles. For these analyses, genotypic data were transformed into pseudodominant "allele phenotypes" (Becher et al. 2000; Rodzen and May 2002) by scoring an allele as present or absent, regardless of dosage. This presence–absence matrix combined all of the alleles at the 10 loci to represent 204 pseudodominant markers.

Genetic diversity and population structure

Green sturgeon is a functional tetraploid (Ludwig et al. 2001) and contains 249 ± 8 chromosomes (Van Eenennaam

Table 4. Pairwise F_{ST} estimates for spawning population and summer aggregation collections used in this study.

	SAC 2002	**SAC 2003**	**SAC 2004**	**SAC 2005**	**SAC 2006**	SPB 2004	SPB 2001	**KLA 1998**	**KLA 2001**	**KLA 2003**	**ROG 2002**
SAC 2002		0.008	0.023	**0.013**	**0.013**	**0.011**	0.006	**0.059**	**0.043**	**0.054**	**0.044**
SAC 2003	**0.007**		**0.015**	**0.006**	**0.010**	**0.002**	0.003	**0.067**	**0.046**	**0.059**	**0.053**
SAC 2004	0.015	**0.012**		**0.022**	**0.015**	**0.018**	**0.014**	**0.085**	**0.067**	**0.078**	**0.071**
SAC 2005	**0.011**	**0.004**	**0.017**		**0.010**	**0.003**	**0.006**	**0.075**	**0.052**	**0.069**	**0.059**
SAC 2006	**0.014**	**0.012**	**0.016**	**0.014**		**0.008**	**0.006**	**0.071**	**0.053**	**0.066**	**0.054**
SPB 2004	**0.009**	0.001	**0.014**	**0.002**	**0.011**		0.001	**0.070**	**0.046**	**0.062**	**0.053**
SPB 2001	**0.003**	0.003	**0.010**	**0.005**	**0.008**	0.001		**0.062**	**0.039**	**0.057**	**0.045**
KLA 1998	**0.053**	**0.074**	**0.080**	**0.079**	**0.072**	**0.075**	**0.064**		**0.015**	**0.009**	**0.012**
KLA 2001	**0.045**	**0.066**	**0.073**	**0.069**	**0.062**	**0.063**	**0.053**	**0.011**		**0.011**	**0.011**
KLA 2003	**0.057**	**0.081**	**0.084**	**0.088**	**0.079**	**0.080**	**0.071**	**0.010**	**0.011**		**0.005**
ROG 2002	**0.043**	**0.065**	**0.071**	**0.069**	**0.061**	**0.064**	**0.053**	**0.010**	**0.009**	**0.006**	
ROG 2000	**0.057**	**0.080**	**0.082**	**0.089**	**0.075**	**0.081**	**0.066**	0.004	**0.014**	**0.015**	**0.010**
ROG 2003	0.056	**0.072**	**0.079**	**0.084**	**0.077**	**0.077**	**0.063**	**0.016**	0.024	**0.013**	0.007
WIN 2002	0.031	**0.056**	**0.060**	**0.056**	**0.047**	**0.053**	**0.041**	**0.008**	**0.007**	**0.008**	0.001
WIN 2000	0.005	**0.011**	**0.019**	**0.014**	**0.015**	**0.010**	**0.006**	**0.037**	**0.029**	**0.038**	**0.025**
COL 1995	**0.011**	**0.008**	0.017	**0.008**	**0.008**	**0.004**	**0.003**	**0.055**	**0.050**	**0.068**	**0.050**
COL 2004	0.004	0.005	0.012	**0.007**	**0.016**	**0.005**	0.002	**0.054**	**0.048**	**0.059**	**0.047**
COL 1999	0.008	**0.009**	**0.020**	**0.012**	**0.014**	**0.009**	**0.010**	**0.045**	**0.038**	**0.046**	**0.037**
WIL 2003	0.005	**0.009**	0.010	**0.011**	**0.011**	**0.009**	**0.004**	**0.048**	**0.043**	**0.055**	**0.039**
GHR 2005	0.018	**0.027**	**0.030**	**0.033**	**0.029**	**0.026**	**0.021**	**0.022**	**0.020**	**0.023**	**0.015**

Note: Above the diagonal are F_{ST} estimates for only the eight tetrasomic loci, while below is the entire 10-loci data set. Bold values indicate significant

et al. 1999). Based on inheritance data (J.A. Israel, unpublished data), we assumed that tetrasomic and disomic loci in green sturgeon were under random assortment of chromosomes from a quadrivalent or bivalent parent without double reduction (Bever and Felber 1992) in all loci except *AfuG41*, *An76*, and *AfuG135*. In these loci, double reduction may be caused by sister chromatids segregating into the same gamete resulting in heterozygote deficiency, which could lead to miscalculation of observed and expected heterozygosity. Observed and Hardy–Weinberg expected heterozygosities at the five tetrasomic loci showed no evidence of double reduction and the two disomic loci were estimated with the programs AUTOTET (Thrall and Young 2000) and GDA (Lewis and Zaykin 2001), respectively. Expected heterozygosity for the tetrasomic loci appearing to undergo double reduction was computed in AUTOTET assuming random chromatid separation with maximum double reduction ($\alpha = 0.142$). To consider the utility of the codominant markers as dominant markers, we calculated the average number of alleles per individual at a locus for the two types of tetrasomic loci in AUTOTET. The number of alleles observed in a collection was calculated using AUTOTET and FSTAT (Goudet 2001).

Pairwise F_{ST} values between collections were calculated in FSTAT to evaluate possible differentiation with 1000 bootstrapped permutations completed to assess significance (Weir and Cockerham 1984). The effect of potentially different rates of mutation and genetic drift among the disomic and tetrasomic markers was considered by calculating F_{ST} values separately for three data sets. The first contained the eight tetrasomic loci, which were transformed by doubling the number of individuals in each collection with pseudogenotypes containing the second half of alleles at these four-dose loci. The second data set contained only the two disomic loci; the third contained all 10 loci. For the 10-loci data set, half of the disomic data were scored as missing, so an equivalent number of individuals could be included for the transformed tetrasomic loci data. These approaches permitted all alleles at a locus to be included in calculating allele frequencies at both tetrasomic and disomic loci. The two data sets including the three loci undergoing double reduction violated assumptions in the calculation of expected heterozygosities for each collection, which may have influenced pairwise F_{ST} calculation.

A neighbor-joining dendrogram of the Cavalli-Sforza–Edwards chord measure (Cavalli-Sforza and Edwards 1967) was created with PHYLIP 3.66 (Felsenstein 2005) with each data set to graphically represent the genetic relationships between all collections. To determine the significance of the neighbor-joining tree's nodes, 1000 bootstrap replicates were completed. The potential number of subpopulations in the entire data set was evaluated with 10 iterations of *structure* 2.2 (Pritchard et al. 2000) with K, the number of populations, between 1 and 10 using the no-admixture model with uncorrelated allele frequencies. The PLOIDY parameter was set to 4 and the two disomic loci scored with two missing alleles each, while burn-in and run lengths of 100 000 and 200 000, respectively, provided Markov chain convergence and consistency in α and q values.

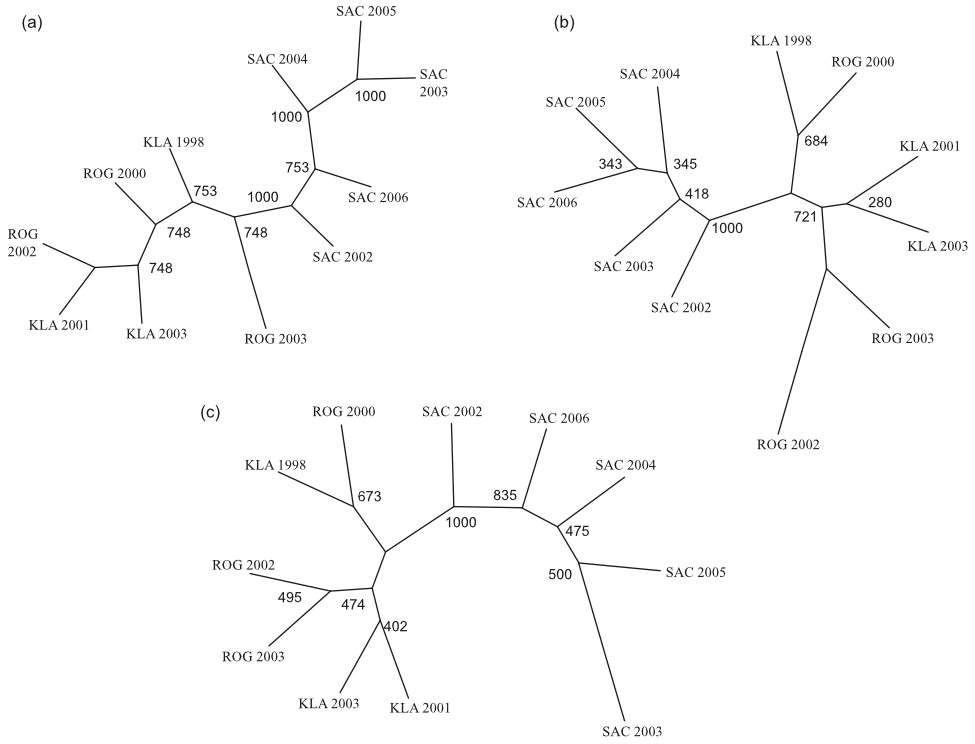
Estimation of DPS contributions to estuarine aggregations

The proportion of fish from each green sturgeon DPS in the estuarine aggregations was determined using three different approaches. First, we applied the likelihood-based mixed stock analysis program *gsi_sim* (Anderson et al. 2008), which assumes a codominant phenotype model. Second, we used the program AFLPOP (Duchesne and Bernatchez 2002), a likelihood-based method for population assignment, which uses a dominant model of allelic expression. Third, we used the Bayesian clustering method of

ROG 2000	ROG 2003	WIN 2002	WIN 2000	COL 1995	COL 2004	COL 1999	WIL 2003	GHR 2005
0.059	**0.063**	**0.028**	**0.011**	**0.013**	**0.007**	0.014	0.009	**0.024**
0.066	**0.062**	**0.035**	**0.010**	**0.008**	**0.004**	**0.008**	**0.011**	**0.026**
0.081	**0.084**	**0.052**	**0.025**	0.018	0.016	**0.025**	0.014	**0.033**
0.078	**0.079**	**0.039**	**0.014**	**0.007**	**0.007**	**0.010**	**0.011**	**0.031**
0.068	**0.073**	**0.034**	**0.013**	**0.010**	**0.013**	**0.015**	**0.010**	**0.029**
0.069	**0.072**	**0.036**	**0.009**	**0.004**	**0.006**	**0.008**	**0.010**	**0.025**
0.058	**0.060**	**0.026**	**0.006**	**0.002**	**0.004**	**0.011**	**0.005**	**0.020**
0.006	**0.018**	**0.010**	**0.038**	**0.050**	**0.053**	**0.046**	**0.044**	**0.021**
0.018	**0.027**	**0.010**	**0.022**	**0.036**	**0.037**	**0.032**	**0.031**	**0.017**
0.018	**0.014**	**0.009**	**0.031**	**0.050**	**0.047**	**0.037**	**0.039**	**0.015**
0.014	0.008	0.001	**0.021**	**0.038**	**0.041**	**0.034**	**0.031**	**0.014**
	0.023	0.008	**0.037**	**0.051**	**0.055**	**0.050**	**0.048**	**0.026**
0.019		**0.016**	**0.034**	**0.054**	**0.048**	**0.045**	**0.043**	**0.018**
0.007	0.015		0.009	**0.019**	**0.023**	0.019	**0.015**	0.006
0.039	**0.034**	0.018		**0.002**	**0.006**	**0.007**	**0.004**	**0.008**
0.063	**0.063**	**0.034**	**0.005**		**0.003**	**0.009**	0.001	**0.015**
0.060	**0.051**	**0.035**	0.005	**0.005**		0.007	0.003	**0.014**
0.051	**0.045**	0.027	0.005	**0.009**	0.007		**0.009**	**0.015**
0.056	**0.049**	**0.031**	**0.004**	**0.002**	0.003	**0.006**		**0.011**
0.027	0.019	0.011	**0.008**	**0.017**	**0.015**	**0.012**	**0.011**	

values at the 5% level corrected for multiple comparisons.

Fig. 2. Neighbor joining tree of Cavalli-Sforza–Edwards unbiased minimum genetic distance of spawning river collections used in the study constructed with data from (*a*) two disomic loci, (*b*) eight tetrasomic loci, and (*c*) all 10 loci. Node support corresponds to proportion of 1000 bootstrapped permutations yielding similar topologies. Collection codes are listed in Table 1.



structure 2.2. In the two likelihood-based analyses, fish of known origin were included as baseline data (also known as "reference individuals"). The Southern DPS was represented in the baseline by pooling the 5 years of fry samples in the Sacramento River (n = 266). The Northern DPS baseline included samples from two populations: 125 adults taken over 3 years from the Klamath represented one source population and 113 adults sampled over 3 years in the Rogue River represented the second. The fraction of fish estimated from the Northern DPS was the sum of the fraction estimated from

Fig. 3. Plots of expected (true, *x*-axis) and observed (model output, *y*-axis) proportions of Southern DPS demonstrate mixture estimation accuracy. (*a*) Each point represents an estimate obtained with one of 5000 training and test set pairs. Training and test sets were obtained by sampling from baseline samples without replacement. (*b*) Number of Southern DPS fish in a test sample ($n = 50$) on the *x*-axis. This displays the variance due solely to inaccuracy in genetic assignment, while Fig. 3*a* includes variance due to binomial sampling of mixture proportion and a finite test sample size of 50.

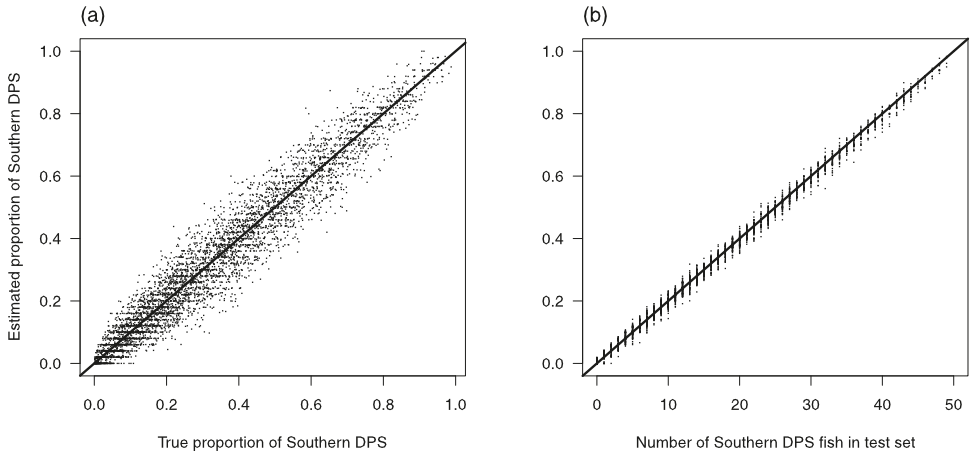
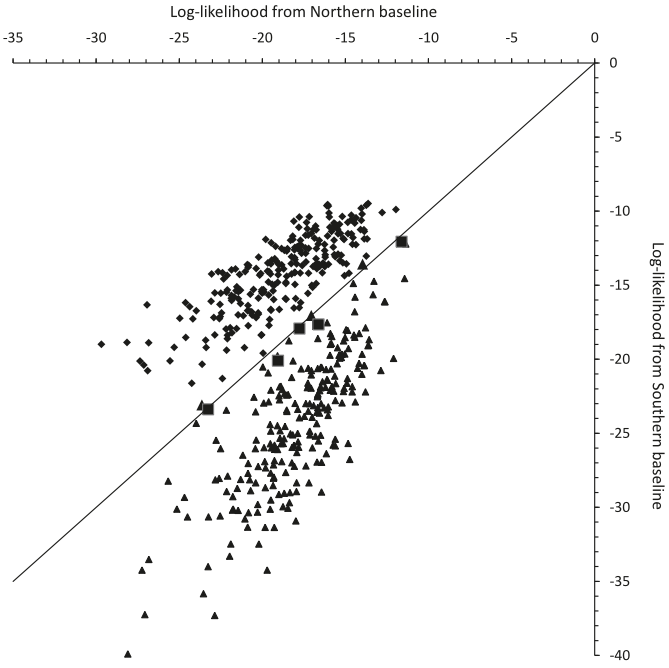


Fig. 4. Log-likelihood values from self-assigned baseline individuals with AFLPOP showing correctly and "misassigned" allelic phenotypes. Small triangles represent correct Northern DPS reallocation and diamonds represent correct Southern DPS reallocations. Large triangles represent incorrectly reallocated Northern DPS individuals and squares represent incorrectly reallocated Southern DPS individuals. The line represents values of equal probability.

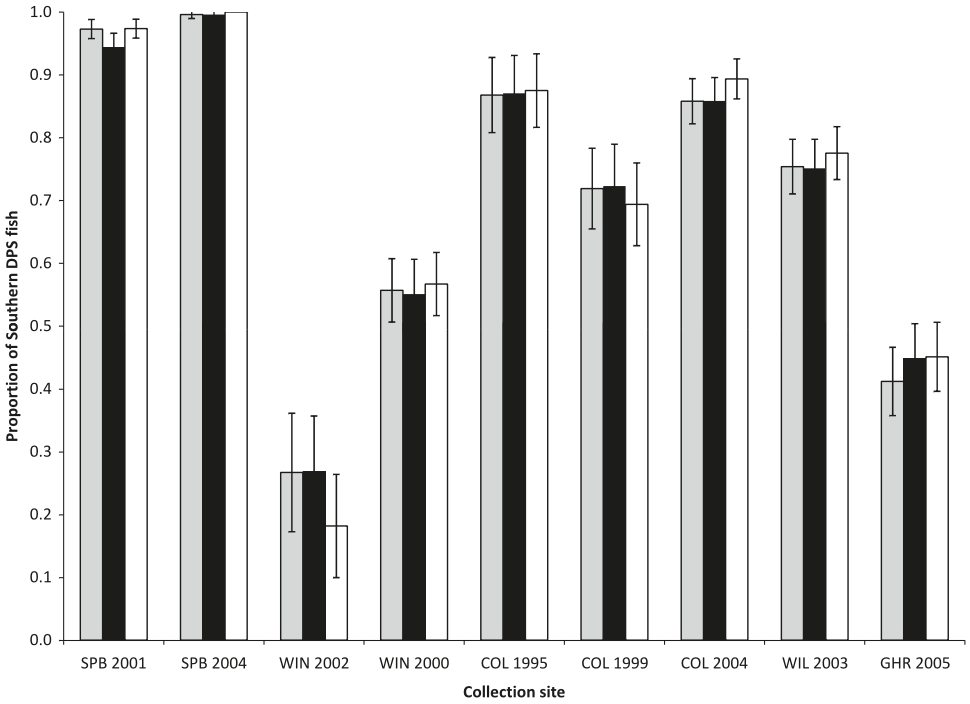


discrete Klamath and Rogue baselines. Reference samples were not used with *structure* because that feature is not available with polyploid data.

The method implemented in *gsi_sim* extends the likelihood model of GMA (Kalinowski 2003) to polysomic data by assuming that the allelic types of all four gene copies are known at each tetrasomic locus. This requires assuming that allelic dosages can be inferred from the electrophoretic data. In the model, the probability of a fish's genotype at any locus follows that of a sample of four from a compound Dirichlet multinomial distribution with parameters determined by the frequency of different alleles in the baselines and a Dirichlet prior of 1/(number of alleles) for each allele in each locus (Rannala and Mountain 1997). The loci are assumed to be in linkage equilibrium, so the multilocus genotype probability is just the product over loci of the single-locus probabilities. The expectation–maximization algorithm (Dempster et al. 1977) is used to find the maximum likelihood estimate of the mixture proportions in an estuary. The estimates of the mixture proportions can then be combined, using Bayes' rule, with a fish's genotype probability to estimate the posterior probability that the fish originated from a given source population. The posterior probability that a fish was from the Northern DPS was calculated as the sum of posterior probabilities that it was from the Klamath or Rogue population. Because the *gsi_sim* method assumes a simple model for tetrasomic inheritance and phenotypic expression that might be violated, we assessed the potential for bias with numerous simulations (see below).

The software AFLPOP (Duchesne and Bernatchez 2002) assumes a dominant model for the data: each allele is counted as either observed or unobserved in an individual. Hence, it does not require assuming that the allelic types of all four gene copies are known or that they are in Hardy–Weinberg equilibrium, alleviating potential problems associated with gene dosage scoring and double reduction. This method assumes that each population's dominant band frequencies are accurate and there is no significant disequilibrium among loci. AFLPOP's likelihood-based method was applied to two different data sets, one with just the eight tetrasomic loci and another with all 10 loci. Because the number of reference individuals allowed in AFLPOP is limited to 254, 12 reference individuals were randomly removed from the 266-fish Southern DPS baseline. The codominant microsatellite data were transformed into pseudodominant allele phenotypes by recording each allele at every locus as present (1) or absent (0), creating a presence–absence matrix

Fig. 5. Proportion of Southern DPS fish in collections of estuarine green sturgeon. Codes as given in Table 1. Standard errors are represented by bars. Shaded bars are *gsi_sim* proportions, solid columns are *structure* proportions, and open bars are AFLPOP proportions.



in which the rows represent alleles and the columns individuals. In the absence of a dominant band in a candidate population, AFLPOP's default frequency correction value of $1/(n + 1)$ was used, where n is the number of samples. AFLPOP computes the log-likelihood of an unknown individual's allele phenotype in each baseline population based on the frequency of the dominant (present) bands at each locus using a variation of the method of Paetkau et al. (1995) and then allocates it into the candidate population with the greatest log-likelihood value.

Like *gsi_sim*, the Bayesian cluster algorithm in *structure* assumes that the allelic types of all four gene copies are known and that population allele frequencies are in Hardy–Weinberg equilibrium with minimal linkage disequilibrium. Results from 10 iterations of *structure* with $K = 2$ were averaged and individuals assigned to the DPS from which they had the highest proportion of ancestry ($q > 0.5$) to obtain the fraction of a collection originating from either DPS.

Assessment of GSI accuracy

To evaluate *gsi_sim*'s performance at estimating mixture proportions and assigning individuals to each DPS, two approaches were taken. The first assessed the self-assignment accuracy of individuals in the baseline (using a leave-one-out procedure). In the second method, we used a cross-validation scheme (Efron and Gong 1983), creating multiple "training" and "test" data sets by sampling individual fish from the baseline randomly and without replacement. The true baseline was sampled with replacement to create a training set containing 216 individuals from the Sacramento collections, 75 individuals from the Klamath collections, and 63 from the Rogue collections. This was done to allow for a consistent baseline sample size and to reserve enough individuals from the baselines to be randomly sampled without replacement into test sets of simulated mixtures ($n = 50$). Training and test sets were replicated 5000 times with mixing proportions of the three baseline populations drawn from a uniform Dirichlet distribution. The proportion of Northern and Southern DPS fish in each simulated mixture was estimated using the training set as the reference sample.

Allocation success rate in AFLPOP was evaluated by removing individual allele phenotypes from the known baseline collections and reallocating them to the baseline population with the highest log-likelihood value. This method of reallocation was compared with AFLPOP's procedure of simulating allelic phenotypes from the allele frequencies of the baseline populations and calculating their log-likelihood values of origin to the baseline populations to assess bias and incorrect assignment.

The accuracy of the assignments made with *structure* was evaluated using the q values for individuals of known origin (Manel et al. 2002; Latch et al. 2006). Since we used no reference sets in the *structure* runs, we first had to decide which of the two clusters corresponded to each DPS. This decision was easily made based on the average q value of reference individuals. We then quantified the rates of correct and incorrect misassignment by designating known-origin individuals with $q > 0.5$ as being correctly assigned and those with $q < 0.5$ as incorrectly assigned.

Identifying cryptic populations in mixed aggregations

The possibility that green sturgeon in estuarine collections may have originated from populations not included in the baseline was evaluated using individuals' negative log-likelihood values from *gsi_sim*. Empirical cumulative density functions (ECDFs) of values from unknown individuals assigned to the Sacramento, Klamath, or Rogue River baselines were compared with the ECDFs of values from

reference individuals, computed using the leave-one-out procedure from the Sacramento, Klamath, and Rogue baselines. Data for multiple years were combined into composite ECDFs for each geographically sampled location using fish that had no more than one locus of missing data. These ECDFs from the unknown and reference individuals were plotted and compared visually and tested for differences using a Kolmogorov–Smirnov (K–S) two-sample test (Sokal and Rohlf 1995).

Results

Genetic diversity and population structure

A total of 1197 green sturgeon samples were genotyped in 20 collections from eight locations (Table 1). Observed heterozygosity among these 20 collections ranged from 0.41 to 0.84 for disomic loci, from 0.22 to 0.79 for doubly reduced tetrasomic loci, and from 0.06 to 0.88 for randomly inherited tetrasomic loci (Table 3). The average number of alleles per individual at a locus ranged between 1.5 and 2.9 for the doubly reduced tetrasomic loci and 1.1 and 3.1 for the randomly inherited tetrasomic loci.

Moderate levels of genetic differentiation were observed among geographically distinct collections of green sturgeon, while low levels of statistically significant temporal differentiation seemed to exist between annual collections from spawning populations regardless of the number of loci in the analyzed dataset. Pairwise F_{ST} values from the 10-locus data set indicated low and moderate levels of genetic differentiation ranging from 0.001 to 0.085 (Table 4). The largest pairwise F_{ST} values (0.043–0.085) were between spawning populations from the Southern and Northern DPSs and almost all inter-DPS pairwise values were significant (85 out of 90). Neighbor-joining dendrograms of the spawning collections demonstrated strong support for the distinctiveness of the DPSs but not for temporal structuring within populations or between the Rogue and Klamath rivers in the Northern DPS, regardless of the number of loci in the analyzed data set (Fig. 2).

The clustering method in *structure* (Pritchard et al. 2000) was run with values of K up to 10 and did not find a minimum likelihood value, although a plot of negative log-likelihood values did appear to reach its asymptote after $K = 8$ (data not shown). Since no value of $\ln \Pr(X|K)$ could be reliably called a maximum, it was not possible to infer a value of K using $\ln \Pr(X|K)$ alone. As K increased to greater than 2, individuals did not show a high fraction of ancestry from any one cluster, and q values for individuals to each inferred cluster became increasingly equal, implying that these solutions for population structure were poor choices (Pritchard et al. 2007). Using 10 runs of *structure* for K between 1 and 10, the mean of the absolute value of the second-order rate of change was calculated and compared with $\ln \Pr(X|K)$ for each K value (Evanno et al. 2005) that showed a single, clear mode for $K = 2$.

Assessment of GSI accuracy with polysomic data

The self-assignment of reference samples showed that individuals from the Southern and Northern DPSs can be distinguished with great accuracy. In *gsi_sim*, 263 of 266 Southern DPS fish (98.8%) were correctly assigned to the Southern DPS. Likewise, 236 (99.2%) of the 238 Northern DPS fish were correctly assigned to the Northern DPS. The self-assignment tests also showed that fish from the two populations within the Northern DPS could not be accurately distinguished: 54.9% and 75.2% of Rogue and Klamath fishes, respectively, were correctly assigned back to their spawning populations. The accuracy of estimates of DPS mixing proportions was confirmed by creating 5000 training and test data sets by permutation (Fig. 3). The fraction of fish from the Southern DPS in the test sets estimated from the Southern DPS are evenly spread about their true value throughout the mixture range from 0 to 1 (Fig. 3*a*), indicating that the mixing proportion can be estimated with little bias. The variance in the estimates due solely to inaccuracy in genetic assignment of individuals within the test set is shown in Fig. 3*b*.

Accuracy in assignment with AFLPOP's dominant expression model where original genotypes from the 10-loci data set were assigned back to one of the two DPSs was also quite high. Using the leave-one-out procedure, 98.7% of the Northern DPS individuals were correctly assigned back to one of the two Northern DPS populations, while 97.2% of the Southern DPS individuals were correctly assigned. Among the two Northern DPS reference groups, 73.2% and 70.6% of Klamath and Rogue individuals, respectively, were correctly assigned back to their original population. The log-likelihood values of individual allelic phenotypes removed and reallocated to the Northern or Southern DPS baseline allelic phenotype frequencies are displayed in Fig. 4. Additionally, AFLPOP simulated 25, 100, and 250 individual allelic phenotypes 100 times from the baseline allelic phenotype frequencies and allocated these phenotypes into the two possible DPSs to test the allocation success rate. Across iterations, simulated allelic phenotypes were correctly assigned to the original DPSs' allelic phenotype frequencies from which they were simulated greater than 99.5% of the time.

Finally, using *structure*, 97.9% of individuals in the Klamath River baseline collection were assigned ($q > 0.5$) to the correct Northern DPS cluster. Likewise, 99.5% of the fish from the Rogue River collection were also assigned to the Northern DPS cluster, and 97.2% of the Sacramento River collection were assigned to the Southern DPS cluster. Ninety-five percent of the reference samples from the Klamath and Rogue rivers had $q \geq 0.98$ estimated to be from the Northern DPS cluster, while 95% of the fish in the Sacramento River baseline had $q \geq 0.94$ from the Southern DPS cluster.

Estimation of estuarine aggregation stock composition

Estuarine aggregations of green sturgeon were determined to originate from both DPSs, although not necessarily reflecting their proximity to spawning populations. The three assignment testing programs yielded very similar estimates of estuarine stock composition (Fig. 5). The Southern DPS was the principal contributor to mixed collections found in San Pablo Bay, the Columbia River, and Willapa Bay. All three methods characterized both of the San Pablo Bay collections as being composed almost exclusively of green sturgeon from the Southern DPS. The estuary collections from the Columbia River and Willapa Bay contained larger pro-

portions (0.69–0.88) of Southern DPS green sturgeon than Northern DPS green sturgeon. The collection from Grays Harbor was estimated to have nearly equal proportions of Northern and Southern DPS green sturgeon, although all three analytical methods estimated slightly more Northern DPS (0.54–0.59) than Southern DPS green sturgeon. The Winchester Bay collections showed the largest range of stock composition variation (0.16–0.55 originating from the Southern DPS) between years and methods, and it could not be explained by the sample size difference alone.

Identifying cryptic populations in mixed aggregations

The negative log-likelihood ECDFs for fish assigned to the Southern DPS from any of the estuaries were all very similar in shape to the ECDF from the Southern DPS baseline (result not shown). In all estuary collections, it was impossible to reject the hypothesis that the distribution of negative log-likelihoods of the Southern DPS assigned fish differed from those of known Southern DPS fish (K–S test, $p > 0.26$ in all cases). The ECDFs of fish in the Willapa and Winchester estuary collections assigned to either the Rogue or Klamath River overlapped well with the ECDFs from the Rogue and Klamath baseline samples, respectively, and were, in fact, statistically indistinguishable (K–S test, $p > 0.21$ in all cases). However, in the Columbia River estuary sample, Rogue-assigned fish had negative log-likelihood values that differed significantly from the Rogue baseline (K–S test, $p = 0.049$), as did their Klamath-assigned counterparts (K–S test, $p = 0.009$). Similarly, in the collections from Grays Harbor, the ECDF of Rogue-assigned fish differed from the baseline (K–S test, $p = 0.01$), as did the ECDF of the Klamath-assigned fish (K–S test, $p = 0.005$). These results could indicate that in the two northernmost estuaries sampled, there were Northern DPS contributions from populations genetically distinguishable from either the Rogue or the Klamath. However, such a conclusion would merit caution, as none of the K–S tests remain statistically significant at the 0.05 level following Bonferroni correction for multiple tests.

Discussion

Comparison of methodological approaches

Polysomic microsatellites have been used in the literature for population structure analyses but have been possibly underutilized in assignment testing due to their perceived problems. The observed heterozygosity and number of alleles per individual per locus were similar among the microsatellite loci regardless of their gene dosage and inheritance pattern, suggesting that combining these data for pairwise population comparisons is a reasonable procedure that does not bias population structure analysis. Observed heterozygosities are likely influenced by relatedness among individuals in spawning collections and admixture of spawning populations in some estuarine collections. Since the number of loci inherited either disomically or tetrasomically under double reduction or independently is restricted, any calculations of linkage disequilibrium and Hardy–Weinberg proportions are similarly limited.

The highly variable nature of the polysomic microsatellite data provided an opportunity for comparing the relative utility of dominant and codominant data for individual assignment. The polymorphism at individual loci was great enough to permit algorithms using frequency data to characterize population identity as Northern and Southern DPSs based on this genetic variation. Moderate levels of genetic variation, supported by F_{ST} and neighbor-joining networks, between spawning populations were large enough to effectively assign unknown green sturgeon individuals into a genetic baseline representative of the principal green sturgeon lineages. The assignment testing procedures used demonstrated that the fractions of Southern DPS fish estimated in west coast estuaries were most likely accurate and reliable, despite the uncertainty in genetic scoring and inheritance of these tetrasomic markers. These results are encouraging and suggest that these assignment testing programs may be useful with other polyploid fish taxa of management interest such as sharks and carp (Leggatt and Iwama 2003).

All three individual assignment methods seemed quite accurate with low bias, and allelic phenotype and genotypic frequency data sets yielded very similar results when tested for accurate self-assignment and allocation of unknown individuals. With *gsi_sim*, much of the variance in test set estimates of Southern DPS composition can be attributed to simple binomial sampling error caused by the fact that each test set contains a random, discrete number of Southern DPS fish. This fraction of simulated fishes assigned to baselines with AFLPOP was likely greater than the ones obtained by self-assignment of reference individuals because the phenotypes of the simulated individuals were simulated according to the inference model AFLPOP used, while the phenotypes of the self-assigned individuals from the reference samples are generated according to the correct, but unknown, natural process.

A better understanding of the inheritance of the tetrasomic markers would assist in adequately characterizing the rate of double reduction and describing the expected heterozygosity of these markers. Also, collections characterizing spawning populations constituted individual and multiple brood years in the cases of the Southern and Northern DPSs, respectively. Inheritance and sampling may influence the assignment of unknown fish to DPSs, although the similar observed heterogeneity and polymorphic content among collections and loci support the use of these collections and makers for individual assignment testing. It appears that genetic differentiation (allelic variation) among DPSs in the baseline is large enough that any unrecognized misinterpretation of genotype variation (allelic frequency) does not affect our results. This was supported by the high success rate of self-assignment and consensus between the methods used for assigning individuals from highly skewed and equally proportioned mixtures.

Stock composition of aggregations

The persistence of genetic differentiation among spawning populations of North American green sturgeon despite significant coastal movement suggests that migration is not a homogenizing force in this broadly distributed species and spawning fidelity has likely generated moderate genetic differences between Northern and Southern DPSs. Although low genetic differentiation among annual collections from spawning rivers was observed, these were grouped into two

baseline groups reflecting both DPSs for the likelihood-based individual assignment testing, since spawning cohorts are overlapping and intergenerational.

These data do not support the hypothesis that all green sturgeon aggregations are composed of the spawning populations that they are closest to and demonstrate a population-scale pattern of Southern DPS specific green sturgeon presence in Northern DPS estuaries. Except for a very small number of Northern DPS green sturgeon in the most southern summer aggregation, estuaries sampled in this study contained both population segments. Our results found very few Northern DPS green sturgeon in the San Pablo Bay collections, supporting the observation that green sturgeon migrate north preferentially once they enter their coastal migration (Lindley et al. 2008). In Winchester Bay, the proportions of green sturgeon from the two DPSs are distinct between the years sampled, although the small sample size of the 2002 collection does not appear to influence the relatively small proportion of Southern DPS fish in that collection. Southern DPS green sturgeon were present in the Columbia River, Willapa Bay, and Grays Harbor collections in greater proportion than expected based on the proximity of these collections to Northern DPS spawning rivers.

This study's results do not support aggregation mixture proportions reflecting the purported abundance of green sturgeon in spawning rivers. Hundreds of adult green sturgeon in the Klamath River are annually harvested, while considerably fewer adult green sturgeon are observed annually in the Sacramento River spawning areas (Adams et al. 2007; M. Gingras, California Department of Fish and Game, Bay Delta Region, 4001 North Wilson Way, Stockton, CA 95205, USA, personal communication (2008); J.A. Israel, unpublished data). Given these limited data, subadult and adult Northern DPS green sturgeon should be more frequently encountered in Northern DPS estuaries than Southern DPS green sturgeon. However, Southern DPS green sturgeon were observed in greater numbers in the majority of estuaries sampled within and outside their DPS boundary and formed the majority of aggregations in the Columbia River, and Willapa Bay and in one Winchester Bay collection. Winchester Bay is the first estuary north of Northern DPS spawning rivers and the variable composition of aggregations may possibly be influenced by nonsynchronous DPS subadult abundances, although no data exist to explore this hypothesis. This study's results suggest that neither spawning river proximity nor adult abundance explains aggregation composition patterns in green sturgeon, and other population-specific processes may contribute to stock composition patterns. It is possible that the southerly sampling on green sturgeon's distribution has biased the apparent dissimilarity between presumed abundance and proportion of occurrence in estuarine collections. Additional collection of green sturgeon from northern Pacific rivers and estuaries would allow for a closer examination of Northern DPS fishes' presence in estuaries.

Green sturgeon migration patterns appear to not be limited to spawning river fidelity, and they may be similar to other sturgeon that use known locations to enhance their fitness and chances for survival (Bemis and Kynard 1997). Southern DPS green sturgeon may demonstrate preferences for estuarine habitat distinct from Northern DPS fishes, and aggregation into preferred estuarine habitats has been demonstrated in other anadromous and marine fishes (Kallio-Nyberg et al. 1999; Corten 2002). The occupation of estuaries by Southern DPS green sturgeon far from their spawning river is similar to the pattern of estuarine occupancy observed for anadromous Gulf sturgeon (*Acipenser oxyrinchus desotoi*) (Sulak and Randall 2002). It is also similar to the Hudson River Atlantic sturgeon's (*Acipenser oxyrinchus oxyrinchus*) propensity for coastal migration into the distant Delaware River and aggregation with sturgeon from other spawning rivers (Waldman et al. 1996). Investigation into mechanisms of habit formation and migratory tradition in green sturgeon and other anadromous sturgeon could provide insight into nongenetic determinants of fidelity and migration behavior.

Efforts to continue interjurisdictional studies to evaluate distribution and abundance of each green sturgeon DPS are critical to the recovery of Southern DPS green sturgeon and the conservation of this species. Recognizing each green sturgeon DPS's complex migratory patterns is necessary for effectively monitoring and minimizing negative impacts on each population segment. Recent closures of estuarine bycatch and sport fisheries involving green sturgeon in California, Oregon, and Washington will benefit the Southern green sturgeon DPS by decreasing the potential for overexploitation of this threatened stock where it occurs in Northern DPS estuaries. In locations such as Winchester Bay and Grays Harbor, where mixing occurs between DPSs, the collection and analysis of additional ecological data on aggregation and capture locations will be useful for identifying each DPS's preferred habitat during estuarine occupancy and migration. Due to stock complexity, the management of green sturgeon will require more focused, intensive interdisciplinary research and monitoring to reduce the uncertainties of alternate restoration and recovery actions for the species.

Acknowledgements

This study was made possible through the collaboration of D. Hillemeier and associates of the Yurok Tribal Fisheries Program, N. Kogut and M. Gingras and associates of the California Department of Fish and Game, B. Poytress and associates of the Red Bluff Fish and Wildlife Service Office, T. Rien, R. Farr, and associates of the Oregon Department of Fish and Wildlife, and O. Langness and associates of the Washington Department of Fish and Wildlife, who all collected and provided samples for this study. Ethan Mora provided the map. Earlier drafts of this manuscript benefited from constructive comments from P. Moyle, B. Ardren, and M. Chilcote as well as three anonymous reviewers. This study has been supported by funds from CALFED (Nos. ERP-02D-P57, 98-C15, and 99-F105), Anadromous Fish Restoration Program (11332-1-G005), Washington Department of Fish and Wildlife, and University of California, Davis, Ecology Graduate Group Block Grants, Hart, Cole, Goss Summer Research Fellowships, and Butler and Walker University Fellowships to J.A.I.

References

Adams, P.B., Grime, C., Hightower, J.E., Lindley, S.T., Moser, M.L., and Parsley, M.J. 2007. Population status of North Amer-

ican green sturgeon, *Acipenser medirostris*. Environ. Biol. Fishes, **79**(3–4): 339–356. doi:10.1007/s10641-006-9062-z.

Allen, P.J., and Cech, J.J., Jr. 2007. Age/size effects on juvenile green sturgeon, *Acipenser medirostris*, oxygen consumption, growth, and osmoregulation in saline environments. Environ. Biol. Fishes, **79**(3–4): 211–229. doi:10.1007/s10641-006-9049-9.

Anderson, E.C., Waples, R.S., and Kalinowski, S.T. 2008. An improved method for predicting the accuracy of genetic stock identification. Can. J. Fish. Aquat. Sci. **65**(7): 1475–1486. doi:10.1139/F08-049.

Beacham, T.D., Gould, A.P., Withler, R.E., Murray, C.B., and Barner, L.W. 1987. Biochemical genetic survey and stock identification of chum salmon (*Oncorhynchus keta*) in British Columbia. Can. J. Fish. Aquat. Sci. **44**(10): 1702–1713. doi:10.1139/f87-209.

Becher, S.A., Steinmetz, K., Weising, K., Boury, S., Peltier, D., Renou, J.-P., Kahl, G., and Wolff, K. 2000. Microsatellites for cultivar identification in *Pelargonium*. Theor. Appl. Genet. **101**(4): 643–651. doi:10.1007/s001220051526.

Bemis, W.E., and Kynard, B. 1997. Sturgeon rivers: an introduction to acipenserform biogeography and life history. Environ. Biol. Fishes, **48**(1–4): 167–183. doi:10.1023/A:1007312524792.

Bever, J.D., and Felber, F. 1992. The theoretical population genetics of autoploidy. *In* Oxford surveys in evolutionary biology. Oxford University Press, Oxford, U.K. pp. 185–217.

Cadrin, S.X., Friedland, K.D., and Waldman, J.R. 2005. Stock identification methods — an overview. *In* Stock identification methods. *Edited by* S.X. Cadrin, K.D. Friedland, and J.R. Waldman. Elsevier, Burlington, Mass. pp. 3–6.

Cavalli-Sforza, L.L., and Edwards, A.W.F. 1967. Phylogenetic analysis. Models and estimation procedures. Am. J. Hum. Genet. **19**(3 Pt. 1): 233–257. PMID:6026583.

Colway, C., and Stevenson, D.E. 2007. Confirmed records of two green sturgeon from the Bering Sea and Gulf of Alaska. Northwest. Nat. **88**(3): 188–192. doi:10.1898/1051-1733(2007)88[188:CROTGS]2.0.CO;2.

Corten, A. 2002. The role of "conservatism" in herring migrations. Rev. Fish Biol. Fish. **11**(4): 339–361. doi:10.1023/A:1021347630813.

De Silva, H.N., Hall, A.J., Rikkerink, E., McNeilage, M.A., and Fraser, L.G. 2005. Estimation of allele frequencies in polyploids under certain patterns of inheritance. Heredity, **95**(4): 327–334. doi:10.1038/sj.hdy.6800728. PMID:16094298.

Dempster, A.P., Laird, N.M., and Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). J. R. Stat. Soc. B, **39**: 1–38.

Duchesne, P., and Bernatchez, L. 2002. AFLPOP: a computer program for simulated and real population allocation, based on AFLP data. Mol. Ecol. Notes, **2**(3): 380–383. doi:10.1046/j.1471-8286.2002.00251.x.

Dugo, M.A., Kreiser, B.R., Ross, S.T., Slack, W.T., Heise, R.J., and Bowen, B.R. 2004. Conservation and management implications of fine-scale genetic structure of Gulf sturgeon in the Pascagoula River, Mississippi. J. Appl. Ichthyol. **20**(4): 243–251. doi:10.1111/j.1439-0426.2004.00572.x.

Efron, B., and Gong, G. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Stat. **37**(1): 36–48. doi:10.2307/2685844.

Erickson, D.L., and Hightower, J.E. 2007. Oceanic distribution and behavior of green sturgeon. *In* Anadromous sturgeons: habitats, threats, and management. *Edited by* J. Munro, D. Hatin, K. McKown, J. Hightower, K.J. Sulak, A.W. Kahnle, and F. Caron. American Fisheries Society Symposium No. 56. American Fisheries Society, Bethesda, Md. pp. 197–211.

Erickson, D.L., and Webb, M.A.H. 2007. Spawning periodicity, spawning migration, and size at maturity of green sturgeon, *Acipenser medirostris*, in the Rogue River, Oregon. Environ. Biol. Fishes, **79**(3–4): 255–268. doi:10.1007/s10641-006-9072-x.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14**(8): 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x. PMID:15969739.

Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6.6. Available from evolution.genetics.washington.edu/phylip.html [accessed 20 September 2006].

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from www2.unil.ch/popgen/softwares/fstat.htm.

Israel, J.A., Cordes, J.F., Blumberg, M.A., and May, B. 2004. Geographic patterns of genetic differentiation among collections of green sturgeon. N. Am. J. Fish. Manag. **24**(3): 922–931. doi:10.1577/M03-085.1.

Kalinowski, S.T. 2003. Genetic Mixture Analysis 1.0. Department of Ecology, Montana State University, Bozeman, MT 59717, USA. Available for download from www.montana.edu/kalinowski.

Kallio-Nyberg, I., Peltonen, H., and Rita, H. 1999. Effects of stock-specific and environmental factors on the feeding migration of Atlantic salmon (*Salmo salar*) in the Baltic Sea. Can. J. Fish. Aquat. Sci. **56**(5): 853–861. doi:10.1139/cjfas-56-5-853.

King, T.L., Lubinski, B.A., and Spidle, A.P. 2001. Microsatellite DNA variation in Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in the Acipenseridae. Conserv. Genet. **2**(2): 103–119. doi:10.1023/A:1011895429669.

Klimley, A.P., Allen, P.J., Israel, J.A., and Kelly, J.T. 2007. Green sturgeon and its environment: past, present, and future. Environ. Biol. Fishes, **79**(3–4): 415–421. doi:10.1007/s10641-006-9177-2.

Lackey, R.T. 2009. Challenges to sustaining diadromous fishes through 2100: lessons learned from western North America. *In* Challenges for diadromous fishes in a dynamic global environment. *Edited by* A. Haro, K.L. Smith, R.A. Rulifson, C.M. Moffitt, R.J. Klauda, M.J. Dadswell, R.A. Cunjak, J.A. Cooper, K.L. Beal, and T.S. Avery. American Fisheries Society Symposium No. 69, American Fisheries Society Symposium, Bethesda, Md. In press.

Latch, E.K., Dharmarajan, G., Glaubitz, J.C., and Rhodes, O.E., Jr. 2006. Relative performance of Bayesian clustering software for inferring population substructuring and individual assignment at low levels of population differentiation. Conserv. Genet. **7**(2): 295–302. doi:10.1007/s10592-005-9098-1.

Leggatt, R.A., and Iwama, G.K. 2003. Occurrence of polyploidy in the fishes. Rev. Fish Biol. Fish. **13**(3): 237–246. doi:10.1023/B:RFBF.0000033049.00668.fe.

Lewis, P.O., and Zaykin, D. 2001. Genetic Data Analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Available from lewis.eeb.uconn.edu/lewishome/software.html [accessed 15 September 2006].

Lindley, S.T., Moser, M.L., Erickson, D.L., Belchik, M., Welch, D.W., Rechisky, E.L., Kelly, J.T., Heublein, J.C., and Klimley, A.P. 2008. Marine migration of North American green sturgeon. Trans. Am. Fish. Soc. **137**(1): 182–194. doi:10.1577/T07-055.1.

Ludwig, A., Belfiore, N.M., Pitra, C., Svirsky, V., and Jenneckens, I. 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics, **158**(3): 1203–1215. PMID:11454768.

Luo, Z.W., Zhang, Z., Zhang, R.M., Pandey, M., Gailing, O., Hattemer, H.H., and Finkeldey, R. 2005. Modeling population ge-

netic data in autotetraploid species. Genetics, **172**(1): 639–646. doi:10.1534/genetics.105.044974. PMID:16172506.

Manel, S., Berthier, P., and Luikart, G. 2002. Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conserv. Biol. **16**(3): 650–659. doi:10.1046/j.1523-1739.2002.00576.x.

Marsden, J.E., Krueger, C.C., and May, B. 1989. Identification of parental origin of naturally produced lake trout in Lake Ontario: application of mixed stock analysis to a second generation. N. Am. J. Fish. Manag. **9**(3): 257–268. doi:10.1577/1548-8675(1989)009<0257:IOPOON>2.3.CO;2.

McQuown, E.C., Sloss, B.L., Sheehan, R.J., Rodzen, J., Tranah, G.J., and May, B. 2000. Microsatellite analysis of genetic variation in sturgeon: new primer sequences for *Scaphirhynchus* and *Acipenser*. Trans. Am. Fish. Soc. **129**(6): 1380–1388. doi:10.1577/1548-8659(2000)129<1380:MAOGVI>2.0.CO;2.

Miller, L.W. 1972. Migrations of sturgeon tagged in the Sacramento – San Joaquin Estuary. Calif. Fish Game, **58**: 102–106.

Moser, M.M., and Lindley, S.T. 2007. Use of Washington estuaries by subadult and adult green sturgeon. Environ. Biol. Fishes, **79**(3–4): 243–253. doi:10.1007/s10641-006-9028-1.

Moyle, P. 2002. Inland fishes of California. University of California Press, Berkeley, Calif.

Paetkau, D., Calvert, W., Stirling, I., and Strobeck, C. 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. **10**: 2375–2388.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**(2): 945–959. PMID:10835412.

Pritchard, J.K., Wen, X., and Falush, D. 2007. Documentation for structure software. Version 2.2. University of Chicago, Chicago, Ill.

Pyatskowit, J.D., Krueger, C.C., Kincaid, H.L., and May, B. 2001. Inheritance of microsatellite loci in the polyploid lake sturgeon (*Acipenser fulvescens*). Genome, **44**(2): 185–191. doi:10.1139/gen-44-2-185. PMID:11341728.

Rannala, B., and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. U.S.A. **94**(17): 9197–9201. doi:10.1073/pnas.94.17.9197. PMID:9256459.

Rodzen, J.A., and May, B. 2002. Inheritance of microsatellite loci in the white sturgeon (*Acipenser transmontanus*). Genome, **45**(6): 1064–1076. doi:10.1139/g02-083. PMID:12502251.

Scribner, K.T., Crane, P.A., Spearman, W.J., and Seeb, L.W. 1998. DNA and allozyme markers provide concordant estimates of population differentiation: analyses of U.S. and Canadian populations of Yukon River fall-run chum salmon (*Oncorhynchus keta*). Can. J. Fish. Aquat. Sci. **55**(7): 1748–1758. doi:10.1139/cjfas-55-7-1748.

Seeb, L.W., Habicht, C., Templin, W.D., Tarbox, K.E., Davis, R.Z., Brannian, L.K., and Seeb, J.E. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the *Exxon Valdez* oil spill. Trans. Am. Fish. Soc. **129**(6): 1223–1249. doi:10.1577/1548-8659(2000)129<1223:GDOSSO>2.0.CO;2.

Shmigirilov, A.P., Mednikova, A.A., and Israel, J.A. 2007. Comparison of biology of the Sakhalin sturgeon, Amur sturgeon, and kaluga from the Amur River, Sea of Okhotsk, and Sea of Japan biogeographic Province. Environ. Biol. Fishes, **79**(3–4): 383–395. doi:10.1007/s10641-006-9050-3.

Smouse, P.E., Waples, R.S., and Tworek, J.A. 1990. A genetic mixture analysis for use with incomplete source population data. Can. J. Fish. Aquat. Sci. **47**(3): 620–634. doi:10.1139/f90-070.

Sokal, R.R., and Rohlf, F.T. 1995. Biometry. W.H. Freeman and Company, New York.

Stephenson, R.L. 1999. Stock complexity in fisheries management: a perspective of emerging issues related to population sub-units. Fish. Res. **43**(1–3): 247–249. doi:10.1016/S0165-7836(99)00076-4.

Sulak, K.J., and Randall, M. 2002. Understanding sturgeon life history: enigmas, myths, and insights from scientific studies. J. Appl. Ichthyol **18**(4–6): 519–528. doi:10.1046/j.1439-0426.2002.00413.x.

Thrall, P.H., and Young, A. 2000. AUTOTET: a program for analysis of autotetraploid genotypic data. J. Hered. **91**(4): 348–349. doi:10.1093/jhered/91.4.348. PMID:10912686.

Van Eenennaam, A.L., Murray, J.D., and Medrano, J.F. 1999. Karotype of the American green sturgeon. Trans. Am. Fish. Soc. **128**(1): 175–177. doi:10.1577/1548-8659(1999)128<0175:KOTAGS>2.0.CO;2.

Van Eenennaam, J.P., Linares, J., Doroshov, S.I., Hillemeier, D.C., Willson, T.E., and Nova, A.A. 2006. Reproductive conditions of the Klamath River green sturgeon. Trans. Am. Fish. Soc. **135**(1): 151–163. doi:10.1577/T05-030.1.

Waldman, J.R., Hart, J.T., and Wirgin, I.I. 1996. Stock composition of the New York Bight Atlantic sturgeon fishery based on analysis of mitochondrial DNA. Trans. Am. Fish. Soc. **125**(3): 364–371. doi:10.1577/1548-8659(1996)125<0364:SCOTNY>2.3.CO;2.

Weir, B., and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution, **38**(6): 1358–1370. doi:10.2307/2408641.

Welsh, A.B., Blumberg, M.A., and May, B. 2003. Identification of microsatellite loci in lake sturgeon, *Acipenser fulvescens*, and their variability in green sturgeon, *A. medirostris*. Mol. Ecol. Notes, **3**(1): 47–55. doi:10.1046/j.1471-8286.2003.00346.x.

Wright, J.E., Jr., Johnson, K., Hollister, A., and May, B. 1983. Meiotic models to explain classical linkage, pseudolinkage, and chromosome pairing in tetraploid derivative salmonid genomes. Isozymes: Curr. Top. Biol. Med. Res. **10**: 239–260. PMID:6354984.

Zane, L., Patarnello, T., Ludwig, A., Fontana, F., and Congiu, L. 2002. Isolation and characterization of microsatellites in the Adriatic sturgeon (*Acipenser naccarii*). Mol. Ecol. Notes, **2**(4): 586–588. doi:10.1046/j.1471-8286.2002.00328.x.

Zhu, B., Liao, X., Shao, Z., Rosenthal, H., and Chang, J. 2005. Isolation and characterization of microsatellites in Chinese sturgeon, *Acipenser sinensis*. Mol. Ecol. Notes, **5**(4): 888–892. doi:10.1111/j.1471-8286.2005.01100.x.