Abstract—Nine microsatellite loci isolated in saffron cod (*Eleginus gracilis*) have potential applications for population genetics. Polymerase chain reaction products of samples of *E. gracilis* from northwestern Alaska amplified reliably, produced only one or two microsatellite bands, and had no apparent homozygote excess. A collection of *E. gracilis* sampled in the Gulf of Alaska (GOA) near Kodiak Island did not amplify reliably at one locus, and allele frequency profiles clustered distinctly (with principal component analysis [PCA]) from the northwestern Alaska collection. Northwestern Alaska and GOA *E. gracilis* collections were genetically different (on the basis of a standardized genetic differentiation measure [G'_{ST}]=0.313, chord distance [D_{chord}]=0.078, P<0.0001) and differed in expected average heterozygosities at shared loci (0.859 and 0.689, respectively). We tested the microsatellite primers on other gadid species endemic to the northern Pacific Ocean, Bering Sea, and Arctic Ocean for cross-species amplification. Not all loci amplified reliably in navaga (*E. nawaga*), Pacific tomcod (*Microgadus proximus*), Arctic cod (*Boreogadus saida*), Pacific cod (*Gadus macrocephalus*), or walleye pollock (*G. chalcogrammus*). Reliable loci varied in microsatellite size profiles and produced distinct PCA clusters and accurate genotype assignments that allowed accurate species identification. The identifications support previous morphological and genetically determined systematic classifications and distinguished the geographically separated collections of *E. gracilis*.

Distinction of saffron cod (*Eleginus gracilis*) from several other gadid species by using microsatellite markers

Noel Sme[1]
Sarah Lyon[1]
Michael Canino[2]
Natalia Chernova[3]
Jason O'Bryhim[4]
Stacey Lance[4]
Kenneth Jones[5]
Franz Mueter[1]
Anthony Gharrett (contact author)[1]

Email address for contact author: a.gharrett@alaska.edu

[1] Juneau Center
College of Fisheries and Ocean Sciences
University of Alaska Fairbanks
17101 Point Lena Loop Road
Juneau, Alaska 99801

[2] Alaska Fisheries Science Center
National Marine Fisheries Service, NOAA
7600 Sand Point Way NE
Seattle, Washington 98115

[3] Zoological Institute of the Russian
   Academy of Sciences
Universitetskaja Naberezhnaya 1

[4] Savannah River Ecology Laboratory
University of Georgia
P.O. Drawer E
Aiken, South Carolina 29802

[5] Section of Hematology-Oncology
Department of Pediatrics
University of Colorado School of Medicine
12800 East 19th Ave, RC-1 North, Room 4129
Aurora, Colorado 80045
St. Petersburg, Russia 199034

The saffron cod (*Eleginus gracilis*) is a gadid fish distributed from the northern Gulf of Alaska (GOA), around the Pacific Rim into the Sea of Okhotsk, and into the Arctic Ocean abutting the North Pacific Ocean (Cohen et al., 1990; Mecklenburg et al., 2016). Mature fish, which generally exceed 20 cm in fork length (FL) and may grow to more than 50 cm FL, are eaten by indigenous Alaskans and in Asia and have potential for commercial harvest in North America (Cohen et al., 1990; NPFMC[1]; Love et al.[2]).

Saffron cod is also an important component of the Arctic ecosystem (Wolotira, 1985; Copeman et al., 2016; Love et al.[2]) and is a significant prey item for several marine mammals (Bluhm and Gradinger, 2008). It is thought to compete for food with Arctic cod (*Boreogadus saida*) and may have a competitive advantage as sea ice changes occur in response to climate change (Love et al.[2]). The species, especially in North Ameri-

[1] NPFMC (North Pacific Fisheries Management Council). 2009. Fishery management plan for fish resources of the Arctic management area, 76 p. NPFMC, Anchorage, AK. [Available from website.]

[2] Love, M. S., N. Elder, C. W. Mecklen-

burg, L. K. Thorsteinson, and T. A. Mecklenburg. 2016. Alaska Arctic marine fish species accounts: saffron cod (*Eleginus gracilis*). *In* Alaska Arctic marine fish ecology catalog. U.S. Geological Survey Sci. Invest. Rep. 2016–5038 (OCS Study, BOEM 2016-048) (L. K. Thorsteinson and M. S. Love, eds.), p. 201–208. [Available from website.]

can waters, is little studied, but its position in the food web, potential population responses to warming and reduction of sea ice in the Arctic, and proposed offshore oil and gas development make learning about this species imperative.

The distributions of several other gadid species— Arctic cod (*B. saida*), Pacific cod (*Gadus macrocephalus*), walleye pollock (*Gadus chalcogrammus*), and Pacific tomcod (*Microgadus proximus*)—overlap with that of *E. gracilis*, and furthermore navaga (*Eleginus nawaga*) from the western Arctic Ocean is a congener of *E. gracilis*. Small gadids of several species are very similar morphologically and often present challenges for identification. The morphological bases of gadiform taxonomy, including the subfamily Gadinae to which all of the species in our study belong, have been described (e.g., Schultz and Welander, 1935; Svetovidov, 1948; Cohen, 1989), as have the phylogenetic relationships among gadiform families (e.g., Roa-Varón and Ortí, 2009) and within Gadinae (Teletchea et al., 2006). However, questions remain about the relationships among *E. gracilis*, *E. nawaga*, and *M. proximus* (e.g., Carr et al., 1999; Roa-Varón and Ortí, 2009). Moreover, the modern geographic separation between *E. eleginus* and *E. nawaga*, if any exists, is unknown.

Genetic analyses of a species can provide insight into several facets of its biology, including population structure, life history (e.g., Kamin et al., 2014), and recent demographic history (e.g., Harpending et al., 1998). Information about population structure can be obtained from surveys in different geographic regions and the fish tested for genetic variation. Microsatellite data are beneficial, when compared with other classes of molecular markers, in that they are often highly polymorphic in fish species (DeWoody et al., 2000) and are relatively inexpensive to apply. Consequently, microsatellite markers were isolated from and developed for *E. gracilis*. Here we 1) examine their variability in two *E. gracilis* collections from geographically separated areas; 2) determine their cross-reactivity with other northern Pacific and Arctic ocean gadids and the ability of suites of these loci to accurately distinguish among species; and 3) evaluate differences in the allele profiles among *M. proximus*, *E. nawaga*, and the two collections of *E. gracilis*.

Materials and methods

Samples and DNA isolation

Collections of *E. gracilis* were collected from the Chukchi Sea in 2011 and near Kodiak Island, Alaska, in 2013. Collections of *B. saida* from the Chukchi Sea were made in 2012 and collections of *E. nawaga* were collected from the Barents Sea in 2013. In 2015, *G. chalcogrammus* were collected in the southeast Bering Sea. Two collections of *M. proximus* were obtained, one from Puget Sound, Washington, between 1997 and 1999, and one from Prince William Sound, Alaska, in

2012. Two collections of *G. microcephalus* were collected in 2013, one collected from Puget Sound and the other from Unimak Pass in the northern Gulf of Alaska (see details for all collections in Table 1).

Tissue samples were preserved in a DNA preservative solution (Seutin et al., 1991) or 95% ethanol and stored in the laboratory at −20°C. Total cellular DNA was isolated with Gentra Puregene[3] or Qiagen DNeasy kits (Qiagen, Hilden, Germany) by following the manufacturer's instructions.

Discovery of microsatellites

An Illumina paired-end shotgun library (Illumina, Inc., San Diego, CA) was prepared by shearing 1 µg of DNA from a single *E. gracilis* Chukchi Sea individual with a Covaris S220 focused-ultrasonicator (**Covaris, Inc.,** Woburn, MA). The standard protocol for the TruSeq DNA library kit (Illumina, Inc.) and a multiplex identifier adaptor index were used (see e.g., Stoutamore et al., 2012). A HiSeq system (Illumina, Inc.) was used to sequence 100-base pair [bp] paired-end readings. The program PAL_FINDER, vers. 0.02.03 (Castoe et al., 2012) was used to analyze 5 million of the resulting sequences to identify readings that had di-, tri-, tetra-, penta-, and hexanucleotide repeat motifs. The data are archived in the Sequence Read Archive of the National Center for Biotechnology Information under accession number SAMN06333955. Once positive reads were identified, oligonucleotide primers were designed with the program Primer3, vers. 2.0.0 (Koressaar and Remm, 2007; Untergasser et al., 2012). To avoid issues with copy number of primer sequences in the genome, loci for which the primer sequences occurred only once or twice in the 5 million reads were selected. Forty-eight presumed loci from *E. gracilis* that met this criterion were chosen for primer design.

The 48 primer pairs were tested with DNA from 8 *E. gracilis* individuals. The polymerase chain reactions (PCRs) were conducted over two 10°C spans of annealing temperatures (65–55°C or 58–48°C) with touch-down thermal cycling profiles (Don et al., 1991). The results (not presented) were analyzed with GeneMapper, vers. 3.7 (Thermo Fisher Scientific, Waltham, MA). Eighteen primer pairs were then selected for evaluation with larger sample sizes.

Analysis of microsatellites

Target sequences of the 18 primer pairs amplified with a touchdown PCR strategy reduced nontarget bands in the product spectrum (Don et al., 1991). All reactions contained ~1 unit Taq polymerase, 1× PCR buffer (50 mM KCl_2, 10 mM Tris-HCl pH 9.0, 0.1% Triton X-100; Promega Corp., Madison, WI), 0.5 µM deoxyribonucleotide triphosphates, and 0.025 to 0.1 µg DNA

---

[3] Mention of trade names or commercial companies is for identification purposes only and does not imply endorsement by the National Marine Fisheries Service

Table 1

Number of samples (*n*), geographic regions, gear used, collector (when known) and collector's affiliation for collections of 6 gadid species sampled in this study: saffron cod (*Eleginus gracilis*), navaga (*E. nawaga*), Pacific tomcod (*Microgadus proximus*), Pacific cod (*Gadus macrocephalus*), walleye pollock (*G. chalcogrammus*), and Arctic cod (*Boreogadus saida*). Asterisks denote specimens originally identified in the field as saffron cod, but were later re-examined.

Species	Scientific name	*n*	Geographic region	Latitude	Longitude
Saffron cod	*Eleginus gracilis*	30	Chukchi Sea	66.91°N	162.55°W
		41	Gulf of Alaska	57.73°N	152.51°W
Nawaga	*Eleginus nawaga*	81	Barents Sea	69.04°N	57.87°E
Pacific tomcod	*Microgadus proximus*	8	Puget Sound	47.71°N	122.52°W
		15	Prince William Sound*	60.87°N	147.19°W
Pacific cod	*Gadus macrocephalus*	5	Puget Sound	48.40°N	124.41°W
		8	Unimak Pass	54.45°N	164.99°W
Walleye pollock	*Gadus chalcogrammus*	6	SE Bering Sea	55.67°N	163.33°W
Arctic cod	*Boregadus saida*	39	Chukchi Sea	66.90°N	162.59°W
		14	Chukchi Sea*	66.90°N	162.59°W

Species	Date	Gear	Collector	Affiliation
Saffron cod	9/11	jig	A. Whiting	Native Village of Kotzebue
	6/7/2013	rod and reel	E. Munk	NOAA Fisheries
Navaga	7/13	trawl	N. Chernova	Russian Academy of Sciences
Pacific tomcod	3/1997-8/1999	beach seine	M Canino	NOAA Fisheries
	7/12	beach seine	M. Arimitsu	U.S. Geological Survey
Pacific cod	3/13	beach seine	M. Canino	NOAA Fisheries
	3/13	trawl	M. Canino	NOAA Fisheries
Walleye pollock	9/15	trawl	W. Strasburger	NOAA Fisheries
Arctic cod	4/12	jig	A. Whiting	Native Village of Kotzebue
	4/12	jig	A. Whiting	Native Village of Kotzebue

template. Fluorescent primers labeled with an IRDye infrared dye (10 µg/mL; Integrated DNA Technologies, Inc., Coralville, IA) were included in the reactions. The amplification profiles for each locus were the following: denaturation at 95°C for 5 min; 20 touchdown cycles at 95°C for 30 s, annealing temperatures ranging from 62 to 52°C (touchdown) for 30 s (decreased 0.5°C per cycle), and 72°C for 30 s; then 15 cycles of 95°C for 30 s, the lowest annealing temperature (55°C) for 30 s, and 72°C for 30 s, and a final extension at 72°C for 5 minutes.

Approximately 1 µL of amplified PCR product and stop buffer (95% formamide, 0.1% bromophenol blue) was loaded onto a 0.25 mm 6% acrylamide gel (PAGE-PLUS™, Amresco, Solon, OH) and fragments were separated in 1× TBE buffer (0.09 M Tris-Borate, 2 mM EDTA, pH 8) at 1500 V with a 4300 DNA Analyzer (LI-COR, Inc., Lincoln, NE). Electrophoresis times varied from 2 to 3 hours depending on allele sizes of the PCR product. The image of the PCR product was analyzed with SAGA, vers. 3.1 (LI-COR, Inc.) software. Two individuals scored each gel separately and samples that differed in recorded allele size were genotyped a second or third time.

Analysis of data

Two collections of *E. gracilis* (one from the Chukchi Sea and another from near Kodiak Island, Alaska) were examined separately (Table 1). Collections of *B. saida* from the Chukchi Sea were combined for analysis as a single species as were collections of *M. proximus* (Prince William Sound and Puget Sound), and of *G. macrocephalus* (Puget Sound and Unimak Pass) (Table 1).

Allele frequencies and expected unbiased heterozygosities were estimated and genotype frequencies were tested for departures from Hardy–Weinberg expectations with GENEPOP, vers. 4.5.1 (Rousset, 2008). Significance of multiple tests was confirmed with sequential Bonferroni tests (Rice, 1989) and false discovery rate (FDR; Benjamini and Hochberg, 1995) corrections. Genotypes of individuals that produced deviations from Hardy–Weinberg expectations or apparent principal component analysis (PCA) outliers were reconfirmed by additional genotyping.

Two genetic distances that are not strongly influenced by the numbers of alleles at a locus, but that are based on very different algorithms, were estimated. The standardized genetic differentiation measure G'_{ST}

Table 2

Microsatellite properties of northern gadid species of the Pacific Rim and Arctic Ocean for the 9 loci designed for saffron cod (*Eleginus gracilis*) sampled in the Chukchi Sea and Gulf of Alaska (GOA) in 2011 and 2013. Number of samples for each species (n), the numbers of different allele observed (n_a), the range of allele sizes, the mean and standard error of the mean (SE) of allele sizes, expected heterozygosities (H_e), and inbreeding coefficients (F_{is}) are given. An entry of dna means the locus did not reliably amplify. Collections were made in 2013 for navaga (*E. nawaga*), during 1997–1999 for Pacific tomcod (*Microgadus proximus*), in 2013 for Pacific cod (*Gadus macrocephalus*), in 2015 for walleye pollock (*G. chalcogrammus*), and in 2012 for Arctic cod (*Boreogadus saida*).

Locus and species	n	n_a	range	mean	H_e	F_{is}
Elgr07						
Chukchi Sea *E. gracilis*	30	10	127–175	155.7 (1.2)	0.867	−0.016
GOA *E. gracilis*	41	7	151–179	160.6 (0.5)	0.683	−0.054
E. nawaga	81	14	115–183	133.7 (0.7)	0.815	−0.028
M. proximus	22	1	123	123.0 (0.0)	0.000	−
G. macrocephalus	14	2	115 and 131	128.7 (1.1)	0.286	−0.130
G. chalcogrammus	6	2	131 and 135	133.3 (0.6)	0.833	−0.667
B. saida	53	dna	−	−	−	−
Elgr11						
Chukchi Sea *E. gracilis*	30	12	208–272	222.1 (1.7)	0.833	0.043
GOA *E. gracilis*	41	8	204–260	214.0 (1.3)	0.634	−0.117
E. nawaga	81	21	240–336	274.7 (1.4)	0.877	0.043
M. proximus	22	17	248–340	285.8 (3.0)	0.727	0.230[a]
G. macrocephalus	14	18	192–204	202.9 (0.6)	0.286	−0.072
G. chalcogrammus	6	dna	−	−	−	−
B. saida	53	dna	−	−	−	−
Elgr31						
Chukchi Sea *E. gracilis*	30	6	191–211	197.1 (0.7)	0.833	−0.103
GOA *E. gracilis*	41	4	191–203	194.8 (0.5)	0.659	−0.015
E. nawaga	81	11	179–231	204.4 (0.8)	0.864	−0.052
M. proximus	22	14	215–267	240.5 (2)	0.955	−0.027
G. macrocephalus	14	18	223–299	263.3 (3.7)	1.000	−0.034
G. chalcogrammus	6	10	215–267	241.7 (4.8)	1.000	−0.035
B. saida	53	37	223–543	355.6 (7.6)	0.962	0.005
Elgr38						
Chukchi Sea *E. gracilis*	30	9	112–144	127.5 (1.1)	0.867	−0.026
GOA *E. gracilis*	41	dna	−	−	−	−
E. nawaga	81	dna	−	−	−	−
M. proximus	22	6	120–140	127.9 (0.8)	0.727	0.068
G. macrocephalus	14	6	128–160	141.9 (1.9)	0.786	0.037
G. chalcogrammus	6	7	236–276	258.0 (4.4)	0.833	0.039
B. saida	53	37	252–448	348.6 (6.3)	0.566	0.422[c]

Table continued

(Hedrick, 2005), based on ratios of heterozygosities adjusted to account for the amount of genetic variation observed at each locus, was estimated with the software program SMOGD, vers. 1.2.5 (Crawford, 2010). Estimates of chord distances (Cavalli-Sforza and Edwards, 1967), a geometric measure, were made with PHYLIP, vers. 3.6 (Felsenstein, 2005).

Principal component analysis was used to contrast the genetic compositions of species groups (SYTAT, vers. 13 software; SYSTAT Software, Inc., San Jose, CA). Correlation matrix-based PCA standardizes variables so that each variable has a similar scale; it was used to contrast the allelic compositions. Covariance matrix-based PCA applies the observed variances so that the scale of variation is included in the analysis;

it was used to contrast allele-frequency profiles. Loci missing from a collection or a species did not contribute to the PCA score.

Assignment tests (GeneClass2; Piry et al., 2004) were used to evaluate the robustness of the differences among species groups. The tests removed each individual fish from the species groups before assignment. The criterion of Rannala and Mountain (1997) was applied in all tests.

Results

Only genotypes from loci that could be reliably interpreted were analyzed in for each species. Nine loci am-

Table 2 (cont.)

Locus and species	n	n_a	range	mean	H_e	F_{is}
Elgr13						
Chukchi Sea *E. gracilis*	30	12	230–286	251.1 (1.3)	0.867	0.007
GOA *E. gracilis*	41	10	226–286	254.3 (1.4)	0.805	0.006
E. nawaga	81	19	214–286	243.8 (1.3)	0.926	−0.009
M. proximus	22	19	242–338	284.5 (3.7)	0.909	0.040
G. macrocephalus	14	14	250–346	314.7 (4.3)	1.000	−0.093
G. chalcogrammus	6	dna	–	–	–	–
B. saida	53	12	206–318	250.8 (1.0)	0.830	−0.064
Elgr14						
Chukchi Sea *E. gracilis*	30	14	322–378	347.6 (1.7)	0.800	0.101
GOA *E. gracilis*	41	9	330–370	345.7 (1.1)	0.829	−0.007
E. nawaga	81	12	318–362	329.4 (0.7)	0.790	−0.010
M. proximus	22	11	326–370	340.3 (1.6)	0.682	0.217
G. macrocephalus	14	4	314–346	325.6 (0.9)	0.143	0.667[b]
G. chalcogrammus	6	10	330–418	364.7 (7.2)	1.000	−0.035
B. saida	53	19	290–366	325.5 (1.7)	0.811	0.121
Elgr44						
Chukchi Sea *E. gracilis*	30	14	212–264	240.9 (1.7)	0.867	0.057
GOA *E. gracilis*	41	7	228–272	247.1 (1.1)	0.537	0.161[a]
E. nawaga	81	14	216–268	238.4 (1.1)	0.840	0.079
M. proximus	22	dna	–	–	–	–
G. macrocephalus	14	dna	–	–	–	–
G. chalcogrammus	6	dna	–	–	–	–
B. saida	53	dna	–	–	–	–
Elgr45						
Chukchi Sea *E. gracilis*	30	13	205–265	218.8 (1.6)	0.867	0.0085
GOA *E. gracilis*	41	4	209–221	213.0 (0.4)	0.683	0.0145
E. nawaga	81	17	189–269	224.5 (1.2)	0.864	0.0471
M. proximus	22	6	197–217	204.9 (0.8)	0.955	−0.0769
G. macrocephalus	14	dna	–	–	–	–
G. chalcogrammus	6	dna	–	–	–	–
B. saida	53	dna	–	–	–	–
Elgr23						
Chukchi Sea *E. gracilis*	30	15	142–202	170.5 (1.8)	0.933	−0.027
GOA *E. gracilis*	41	4	162–190	168.1 (0.4)	0.683	−0.181
E. nawaga	81	17	138–214	168.1 (1.1)	0.926	−0.019
M. proximus	22	13	138–206	161.6 (2.4)	0.909	−0.044
G. macrocephalus	14	17	154–286	215.0 (5.0)	0.929	0.034
G. chalcogrammus	6	11	186–318	246.7 (12.6)	1.000	−0.017
B. saida	53	23	138–258	191.6 (2.3)	0.660	0.309[c]

[a] $P<0.05$; [b] $P<0.01$; [c] $P<0.001$.

plified reliably and had no apparent homozygote excess in *E. gracilis* (Table 2; Suppl. Table 1). However, not all loci that were reliable in *E. gracilis* amplified consistently and produced just 1 or 2 bands in all sets of samples. Most notably, Elgr38 did not amplify reliably in GOA samples of *E. gracilis*, nor was it reliable in *E. nawaga*. In addition, only 7 of the 9 loci worked well in *M. proximus* and only 5 in either *G. chalcogrammus* or *B. saida*. Again, Elgr38 did not amplify reliably in the GOA samples of *E. gracilis* nor was it reliable in *E. nawaga*. In addition, only 7 of the loci worked well in *M. proximus* and only 5 in each of *G. chalcogrammus* and *B. saida*. Of the loci that did not amplify reliably for a species group, several did produce bands. Only

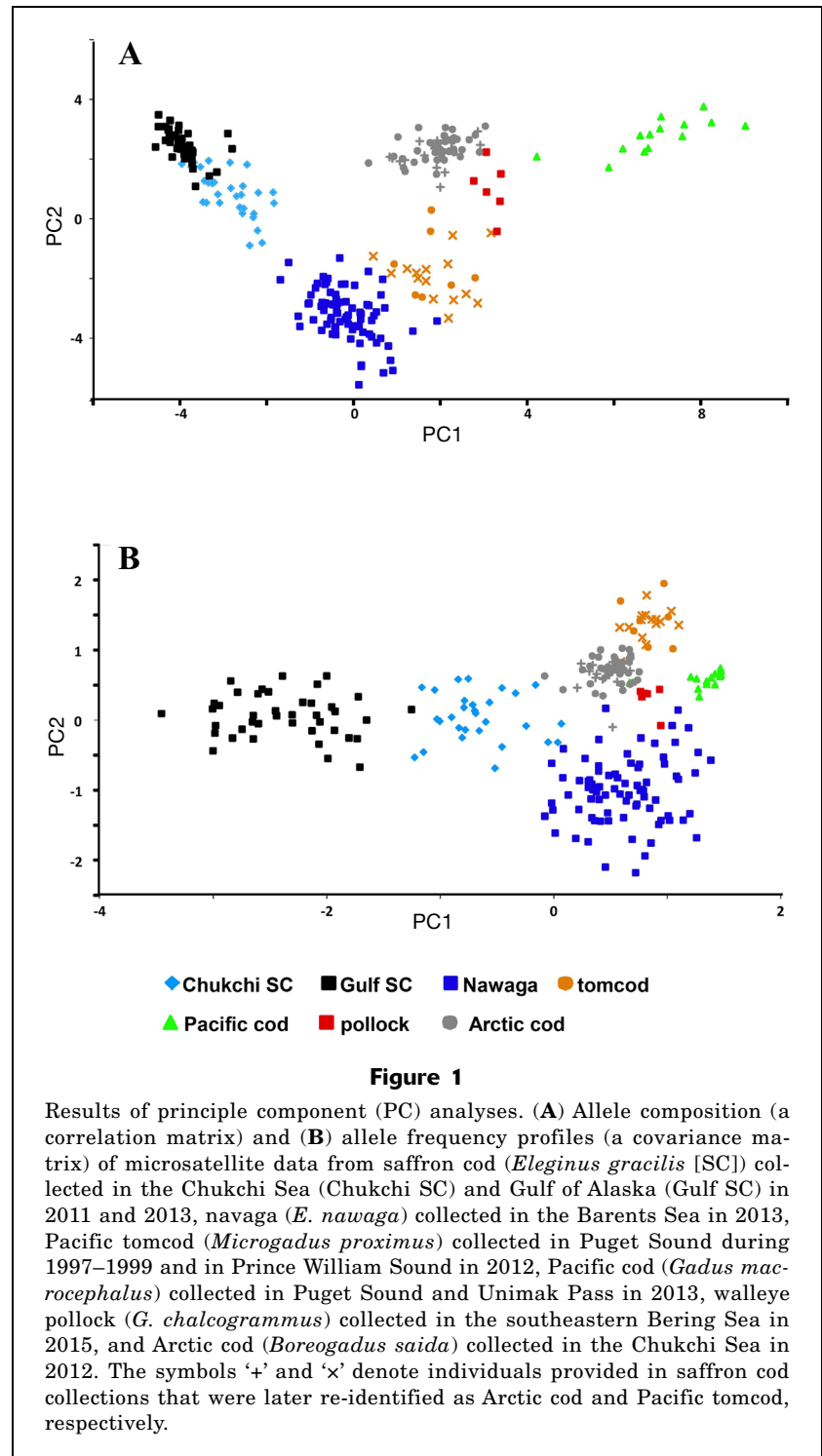the loci that could be interpreted reliably were analyzed in each species.

Comparisons of gadid collections

Differences in ranges of allele sizes differentiated species and species groups (Table 2, Suppl. Fig. 1). For example, alleles at Elgr38 were on average much larger for *B. saida* and *G. chalcogrammus* than for the other species; alleles at Elgr31 were larger on average for *B. saida* and alleles at Elgr23 were on average larger for *G. macrocephalus* and *G. chalcogrammus*. The divergences in allele frequency size ranges were reflected in values of D_{chord} and G'_{ST} (Table 3), all of which were

significant (adjusted pairwise homogeneity tests $P<0.0001$). The estimate of G'_{ST} between the two *E. gracilis* collections was smaller than values between all other gadid pairs; whereas the estimate of D_{chord} was smaller than that of all but three of the comparisons of gadids, even though different suites of microsatellite loci were used. To provide a comparison of the extent of divergence between the two *E. gracilis* collections, values of G'_{ST} and D_{chord} were estimated for the species pair *Sebastes aleutianus* and *S. melanostictus* from data in Gharrett et al. (2005), $G'_{ST}=0.551$ and $D_{chord}=0.064$. The estimate of G'_{ST} between the *E. gracilis* pair was lower (0.313) but the estimate of D_{chord} was higher (0.078) than that between *S. aleutianus* and *S. melanostictus*, presumably because different algorithms were applied; D_{chord} has a geometric basis and G'_{ST} is based on ratios of heterozygosities adjusted to account for the amount of genetic variation observed at each locus (Hedrick, 2005).

Individual-based PCA of allelic compositions (a correlation matrix) and allele frequency profiles (a covariance matrix) produced both species- and collection-specific clusters (Fig. 1). The plot of the first and second components of the correlation-based PCA separated individual species more clearly, but separation of the two *E. gracilis* collections was not as strong. The covariance-based PCA clearly separated the two *E. gracilis* collections, but the other species were not separated quite as well. The first five components of the correlation-based analysis accounted for 10.6% and the first two components accounted for 5.1% of the overall variation in allelic composition. In contrast, the first five components of the covariance-based PCA accounted for 24.3% and the first two for 14.1% of the overall variation in allelic frequencies. Nevertheless, sufficient variation existed to separate these species and the two collections of *E. gracilis*.

A series of 4 tests was needed to estimate assignments of individuals because not all loci could be used for all species groups (Suppl. Table 2). The tests were the following: 1) all individuals were assigned on the basis of the three loci all groups had in common—Elgr14, Elgr23, and Elgr31; 2) the individuals scored in 1) as Chukchi Sea *E. gracilis* (CSC), GOA *E. gracilis* (GSC), *E. nawaga* (NAW), *M. proximus*

Figure 1

Results of principle component (PC) analyses. (**A**) Allele composition (a correlation matrix) and (**B**) allele frequency profiles (a covariance matrix) of microsatellite data from saffron cod (*Eleginus gracilis* [SC]) collected in the Chukchi Sea (Chukchi SC) and Gulf of Alaska (Gulf SC) in 2011 and 2013, navaga (*E. nawaga*) collected in the Barents Sea in 2013, Pacific tomcod (*Microgadus proximus*) collected in Puget Sound during 1997–1999 and in Prince William Sound in 2012, Pacific cod (*Gadus macrocephalus*) collected in Puget Sound and Unimak Pass in 2013, walleye pollock (*G. chalcogrammus*) collected in the southeastern Bering Sea in 2015, and Arctic cod (*Boreogadus saida*) collected in the Chukchi Sea in 2012. The symbols '+' and '×' denote individuals provided in saffron cod collections that were later re-identified as Arctic cod and Pacific tomcod, respectively.

(PTC), and *G. macrocephalus* (PCO) were assigned on the basis of Elgr7, Elgr11, Elgr13, Elgr14, Elgr23, and Elgr31; 3) the individuals scored in 2) as CSC, GSC, or NAW were tested at Elgr7, Elgr11, Elgr13, Elgr14, Elgr23, Elgr31, Elgr44, and Elgr45; and 4) the individuals scored in 1) as PTC, PCO, *G. chalcogrammus* (WPO), or *B. saida* (ACO) were tested at Elgr14,

Table 3

Estimates of pairwise chord distances (D_{chord}; above the diagonal) and standardized genetic differentiation measure (G'_{ST}, below the diagonal) for saffron cod (*Eleginus gracilis*) sampled in the Chukchi Sea and Gulf of Alaska (GOA) in 2011 and 2013 and for navaga (*E. nawaga*), Pacific tomcod (*Microgadus proximus*), Pacific cod (*Gadus macrocephalus*), walleye pollock (*G. chalcogrammus*), and Arctic cod (*Boreogadus saida*) sampled in the Pacific Rim and Arctic Ocean during 1997–2015. All estimates were significant (adjusted probabilities: $P<0.0001$). Values of average unbiased expected heterozygosity (H_e) are indicated in italic type on the diagonal.

Collection	A	B	C	D	E	F	G
A Chukchi Sea *E. gracilis*	*0.859*	0.078[a]	0.076[a]	0.138[a]	0.189[b]	0.218[d]	0.076[d]
B GOA *E. gracilis*	0.313	*0.689*	0.130[a]	0.183[b]	0.245[b]	0.296[e]	0.095[e]
C *E. nawaga*	0.414	0.680	*0.863*	0.093[b]	0.137[c]	0.158[e]	0.069[e]
D *M. proximus*	0.603	0.779	0.565	*0.733*	0.182[b]	0.228[d]	0.088[e]
E *G. macrocephalus*	0.877	0.963	0.822	0.721	*0.633*	0.204[d]	0.092[e]
F *G. chalcogrammus*	0.868	0.893	0.739	0.582	0.449	*0.933*	0.087[e]
G *B. saida*	0.599	0.680	0.584	0.781	0.681	0.607	*0.766*

[a] 8 loci;
[b] 7 loci;
[c] 6 loci;
[d] 5 loci;
[e] 4 loci for both D_{chord} and G'_{ST} estimates.

Elgr23, Elgr31, and Elgr38. The results of 3) and 4) assigned each individual to its own group, except 1 CSC (96.7% of the total) and 1 ACO (98.1% of the total) (Table 4).

Previous molecular studies have recognized *G. macrocephalus*, *G. chalcogrammmus*, and *B. saida* as distinct species (Coulson et al., 2006, Carr et al., 1999) but the systematic relationships among *E. gracilis*, *E. nawaga*, and *M. proximus* are still unresolved (Mecklenburg et al., 2016). Differences in the allele frequency profiles are easier to see in plots that include only those four groups (Table 2, Suppl. Fig. 2). The *M. proximus* and *E. nagawa* distributions clearly differ from those of the 2 *E. gracilis* collections at Elgr07 and Elgr11. The profiles for *M. proximus* and *E. nagawa* clearly differ from those for the 2 collections of *E. gracilis* at Elgr07 and Elgr11. *M. proximus* also differs at Elgr13 and Elgr31 and has a substantially higher number of large alleles. The numbers of observed alleles (Table 2) in the collection of GOA *E. gracilis* are relatively lower than those of the others and several are more abundant (Suppl. Fig. 2), which is consistent with the somewhat lower heterozygosity (Table 2) of the GOA *E. gracilis*.

Discussion

Eight of the nine microsatellites that were evaluated for two collections of *E. gracilis* and that amplified reliably were variable (heterozygosities 0.537 to 0.933) and had no apparent homozygote excess, indicating low null allele frequencies. The single exception, Elgr38, amplified reliably for the Chukchi Sea collection of *E. gracilis* but not for the GOA collection. At the other loci, the two collections had similar allele size ranges but differed substantially in allele frequencies ($G'_{ST}=0.313$, $D_{chord}=0.078$, $P<0.0001$). The observed differences were similar to those between two cryptic rockfish species that had overlapping ranges, *S. aleutianus* and *S. melanostictus*, although they were estimated with different suites of loci. In the PCA plots, individuals from the two collections of *E. gracilis* were mostly distinct from each other, particularly in the analysis of the covariance matrix, which focuses on the allele frequencies rather than allele composition. It is also notable that the PCA analyses included frequency differences of the other gadids analyzed, and differences between the 2 collections of *E. gracilis* were evident against the background variation from other species.

Assignment tests placed all but one saffron cod in the group from which it originated. Not all nine microsatellite loci amplified reliably in all of the other gadid species analyzed and some had an excess of homozygotes, most likely as a consequence of null alleles; those loci were not used for assignment tests. Nevertheless, where comparisons were possible, all the other gadids differed in microsatellite composition ($P<0.0001$) from both collections of *E. gracilis* and each other. The correlation matrix-based PCA, in particular, clustered individuals according to species or geographic groups of species. The PCA analyses turned out to be valuable in analyzing a large set of samples of putative *E. gracilis* because the analysis revealed outliers that, when compared with the clusters of other gadids, enabled detection of individuals misidentified as *E. gracilis*. Two

Table 4

Summary of results of a series of tests (Piry et al., 2004) that assigned each fish to 1 of 7 species groups: saffron cod (*Eleginus gracilis*) of the Chukchi Sea (CSC), saffron cod of the Gulf of Alaska (GSC), navaga (*E. nawaga*) (NAW), Pacific tomcod (*Microgadus proximus*) (PTC), Pacific cod (*Gadus macrocephalus*) (PCO), walleye pollock (*G. chalcogrammus*) (WPO), and Arctic cod (*Boreogadus saida*) (ACO). *n*=the number of individuals of each group. For all results from the assignment tests, see Supplementary Table 2.

		Assigned to						
n	Species group	CSC	GSC	NAW	PTC	PCO	WPO	ACO
30	CSC	29	1[a]	0	0	0	0	0
41	GSC	0	41	0	0	0	0	0
81	NAW	0	0	81	0	0	0	0
23	PTC	0	0	0	23	0	0	0
14	PCO	0	0	0	0	14	0	0
6	WPO	0	0	0	0	0	6	0
53	ACO	0	0	0	0	0	1[b]	52

[a] 83% GSC/ 17% CSC.
[b] 55% PCO/ 44% WPO/ 1% ACO.

notable instances were 14 aberrant genotypes included in a collection of *E. gracilis* from the Chukchi Sea and another 15 in a collection of *E. gracilis* from Prince William Sound. In both instances, it was possible to re-examine the individual specimens; the former were re-identified as *B. saida* and the latter as *M. proximus* (Table 1). Both sets of re-identified individuals were included with their correct species in the analyses presented here (designated as '+' and '×', respectively in Fig. 1). Assignment tests correctly reassigned all of the other gadids except one Arctic cod.

In these analyses, the two collections of *E. gracilis*, and the collections of *M. proximus*, and *E. nawaga* were all distinct from each other (*P*<0.0001). The degree of their divergences mostly exceeded those observed between *S. aleutianus* and *S. melanostictus* (Gharrett et al., 2005) and each of the collections clustered separately in PCAs. It is notable that misidentified individuals of Prince William Sound *M. proximus* were collected at the same site with *E. gracilis*, but were genetically distinct from them. Clearly, some field identifications, even by trained personnel, are challenging (cf. Teletchea, 2009). It is unlikely that they represent two sympatric populations of a single marine species—populations that are so strongly different genetically. Although it could be argued that the genetic differences between the collections of *E. nawaga* and *E. gracilis* could result from divergence over the large distance that separates them, the very large divergences in allele frequencies, as well as similar differences in allele size ranges at Elgr11 and Elgr14, are more consistent with their being distinct species. More complete information on the modern Arctic distributions of the two species of *Eleginus*, and the location of the historic contact zone between them, would contribute to re-

solving their systematic status, as would independent data, such as mitogenomic sequences of *E. nawaga* and *E. gracilis*, coupled with morphological characters (Teletchea, 2009).

Acknowledgements

Literature cited

Benjamini, Y., and Y. Hochberg.
    1995.   Controlling the false discovery rate: a practical and powerful approach to multiple testing.   J. R. Stat. Soc., B. 57:289–300.   Article
Bluhm, B. A., and R. Gradinger.
    2008.   Regional variability in food availability for arctic marine mammals.   Ecol. Appl. 18:S77–S96.   Article
Castoe, T. A., A. W. Poole, A. P. J. de Koning, K. L. Jones, D. F. Tomback, S. J. Oyler-McCance, J. A. Fike, S. L. Lance, J. W. Streicher, E. N. Smith, and D. D. Pollack.
    2012.   Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake.   PLoS ONE 7(2):e30953.   Article

Cavalli-Sforza, L. L, and A. W. F. Edwards.
    1967.   Phylogenetic analysis: models and estimation proce-
        dures.   Evolution 21:550–570.   Article
Crawford, N. G.
    2010.   SMOGD: software for the measurement of genetic
        diversity.   Mol. Ecol. Resour. 10:556–557.
Carr, S. M., and H. D. Marshall.
    2008.   Phylogeographic analysis of complete mtDNA ge-
        nomes from Walleye Pollock (*Gadus chalcogrammus* Pal-
        las, 1811) shows an ancient origin of genetic biodiversity.
        DNA Sequence 19:490–496.   Article
Carr, S. M., D. S. Kivlichan, P. Pepin, and D. C. Crutcher.
    1999.   Molecular systematics of gadid fishes: implications
        for the biogeographic origins of Pacific species.   Can. J.
        Zool. 77:19–26.   Article
Cohen, D. M. (ed.).
    1989.   Papers on the systematics of gadiform fishes.   Nat.
        Hist. Mus. Los Ang. Cty., Sci. Ser. 32, 262 p.
Cohen, D. M., T. Inada, T. Iwamoto, and N. Scialabba.
    1990.   FAO species catalogue, vol. 10. Gadiform fishes of
        the world (Order Gadiformes). An annotated and illus-
        trated catalogue of cods, hakes, grenadiers and other
        gadiform fishes known to date.   FAO Fisheries Synopsis
        125, 442 p.   FAO, Rome.
Copeman, L. A., B. J. Laurel, K. M. Boswell, A. L. Sremba,
    K. Klinck, R. A. Heintz, J. Vollenweider, T. E. Helser, and
    M. L. Spencer.
    2016.   Ontogenetic and spatial variability in trophic bio-
        markers of juvenile saffron cod (*Eleginus gracilis*) from
        the Beaufort, Chukchi and Bering Seas.   Polar Biol.
        39:1109–1126.   Article
Coulson, M. W., H. D. Marshall, P. Pepin, and S. M. Carr.
    2006.   Mitochondrial genomics of gadine fishes: implica-
        tions for taxonomy and biogeographic origins from whole-
        genome data sets.   Genome 49:1115–1130.   Article
DeWoody, J. A., and J. C. Avise.
    2000.   Microsatellite variation in marine, freshwater and
        anadromous fishes compared with other animals.   J.
        Fish Biol. 56:461–473.   Article
Don, R. H., P. T. Cox, B. J. Wainwright, K. Baker, and J. S.
    Mattick.
    1991.   'Touchdown' PCR to circumvent spurious prim-
        ing during gene amplification. Nucleic Acids Res.
        19:4008.   Article
Felsenstein, J.
    2005.   PHYLIP (Phylogeny Inference Package), vers. 3.6.
        Dep. Genome Sci., Univ. Washington, Seattle, WA.
        [Available from website.]
Gharrett, A. J., A. P. Matala, E. L. Peterson, A. K. Gray, and
    Z. Li.
    2005.   Two genetically distinct forms of rougheye rockfish
        are different species.   Trans. Am. Fish. Soc. 134:242–
        260.   Article
Harpending, H. C, M. A. Batzer, M. Gurven, L. B. Jorde, A. R.
    Rogers, and S. T. Sherry.
    1998.   Genetic traces of ancient demography.   Proc. Natl.
        Acad. Sci. U.S.A. 95:1961–1967.   Article
Hedrick, P. W.
    2005.   A standardized genetic differentiation mea-
        sure.   Evolution 59:1633–1638.   Article
Kamin, L. M., K. J. Palof, J. Heifetz, and A. J. Gharrett.
    2014.   Interannual and spatial variation in the population
        genetic composition of young-of-the-year Pacific ocean

perch (*Sebastes alutus*) in the Gulf of Alaska.   Fish.
        Oceanogr. 23:1–17.   Article
Koressaar, T., and M. Remm.
    2007.   Enhancements and modifications of primer design
        program Primer3. Bioinformatics 23:1289–1291.   Article
Mecklenburg, C. W., T. A. Mecklenburg, B. A. Sheiko, and D.
    Steinke.
    2016.   Pacific Arctic marine fishes, 375 p.   Conservation
        of Arctic Flora and Fauna, Akureyri, Iceland.
Piry, S., A. Alapetite, J.-M. Cornuet, D. Paetkau, L. Baudouin,
    and A. Estoup.
    2004.   GENECLASS2: A software for genetic assign-
        ment and first-generation migrant detection.   J. Hered.
        95:536–539.   Article
Rannala, B., and J. L. Mountain.
    1997.   Detecting immigration by using multilocus geno-
        types.   Proc. Natl. Acad. Sci. U.S.A. 94:9197–9201.
        Article
Rice, W. R.
    1989.   Analyzing tables of statistical tests.   Evolution
        43:223–225.   Article
Roa-Varón, A., and G. Ortí.
    2009.   Phylogenetic relationships among families of
        Gadiformes (Teleostei, Paracanthopterygii) based on
        nuclear and mitochondrial data.   Mol. Phylogen. Evol.
        52:688–704.   Article
Rousset, F.
    2008.   GENEPOP'007: a complete re-implementation of
        the GENEPOP software for Windows and Linux.   Mol.
        Ecol. Resour. 8:103–106.   Article
Schultz, L. P., and A. D. Welander.
    1935.   A review of the cods of the northeastern Pacific with
        comparative notes on related species.   Copeia 1935:127–
        139.   Article
Seutin, G., B. N. White, and P. T. Boag.
    1991.   Preservation of avian blood and tissue samples for
        DNA analysis.   Can. J. Zool. 69:82–90.   Article
Stoutamore, J. L., C. N. Love, S. L. Lance, K. L. Jones, and
    D. Tallmon.
    2012.   Development of polymorphic microsatellite mark-
        ers for blue king crab (*Paralithodes platypus*).   Conserv.
        Genet. Resour. 4:897–899.   Article
Svetovidov, A. N.
    1948.   Gadiformes (Treskoobraznye).   Fauna of the U.S.S.R.,
        vol. 9, no. 4, 304 p.   Israel Program Sci. Transl., Jeru-
        salem, Israel.
Teletchea, F.
    2009.   Molecular identification methods of fish species:
        reassessment and possible applications.   Rev. Fish Biol.
        Fish. 19:265–293.   Article
Teletchea, F., V. Laudet, and C. Hännia.
    2006.   Phylogeny of the Gadidae (sensu Svetovidov, 1948)
        based on their morphology and two mitochondrial genes.
        Mol. Phylogen. Evol. 38:189–199.   Article
Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Fair-
    cloth, M. Remm, and S. G. Rozen.
    2012.   Primer3—new capabilities and interfaces.   Nucleic
        Acids Res. 40:e115.   Article
Wolotira, R. J., Jr.
    1985.   Saffron cod *(Eleginus gracilis)* in western Alaska,
        the resource and its potential.   NOAA Tech. Memo.
        NMFS F/NWC-79, 119 p.