Canadian Science Publishing

ARTICLE

Genome-wide analysis reveals regional patterns of drift, structure, and gene flow in longfin smelt (*Spirinchus thaleichthys*) in the northeastern Pacific

İsmail K. Sağlam, James Hobbs, Randall Baxter, Levi S. Lewis, Alyssa Benjamin, and Amanda J. Finger

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Abstract: The southernmost stock of longfin smelt (*Spirinchus thaleichthys*) is approaching extirpation in the San Francisco Estuary (SFE); however, patterns of genetic structure, diversity and gene flow which are vital for management are poorly understood in this species. Here, we use genome-wide data to evaluate population structure of longfin smelt across a broad latitudinal scale across estuaries ranging from the SFE to Yakutat Bay and Lake Washington, and fine scale within the Fraser River and the SFE. Results indicate high genetic structure between major estuaries, fine-scale structure within the Fraser River, and low levels of structure within the SFE. Genetic structure was more pronounced between northern estuaries whereas southern estuaries showed shared ancestry and ongoing gene flow, most notably unidirectional northward migration out of the SFE. Furthermore, we detected signatures of local adaptation within the Fraser River and the Skeena River estuaries. Taken together, our results identify broad patterns of genetic diversity in longfin smelt shaped by co-ancestry, unidirectional migration and local adaptation. Results also suggest that the SFE population is genetically distinct from northernmost populations and an important source for maintaining nearby populations.

Résumé : Si le stock le plus méridional d'éperlans d'hiver (*Spirinchus thaleichthys*) est en voie de disparaître de l'estuaire de San Francisco (ESF), les motifs de structure génétique, de diversité et de flux génétique chez cette espèce, qui sont d'importance vitale pour sa gestion, ne sont pourtant pas bien compris. Nous utilisons des données pangénomiques pour évaluer la structure des populations d'éperlans d'hiver sur une grande fourchette de latitudes dans des estuaires allant de l'ESF à la baie de Yakutat et au lac Washington, et à échelle fine dans le fleuve Fraser et l'ESF. Les résultats indiquent une forte structure génétique entre de grands estuaires, une structure fine dans le fleuve Fraser et une faible structuration dans l'ESF. La structure génétique est plus prononcée entre estuaires septentrionaux, alors que les estuaires méridionaux présentent une ascendance commune et un flux génétique toujours en cours, notamment une migration unidirectionnelle vers le nord à partir de l'ESF. Nous relevons en outre des signatures d'adaptation locale dans les estuaires des fleuves Fraser et Skeena. Collectivement, nos résultats font ressortir de grands motifs de diversité génétiques chez les éperlans d'hiver modelés par la coascendance, la migration unidirectionnelle et l'adaptation locale. Ils indiqueraient également que la population de l'ESF est distincte sur le plan génétique des populations les plus septentrionales et qu'elle constitue une source importante pour le maintien des populations environnantes. [Traduit par la Rédaction]

Introduction

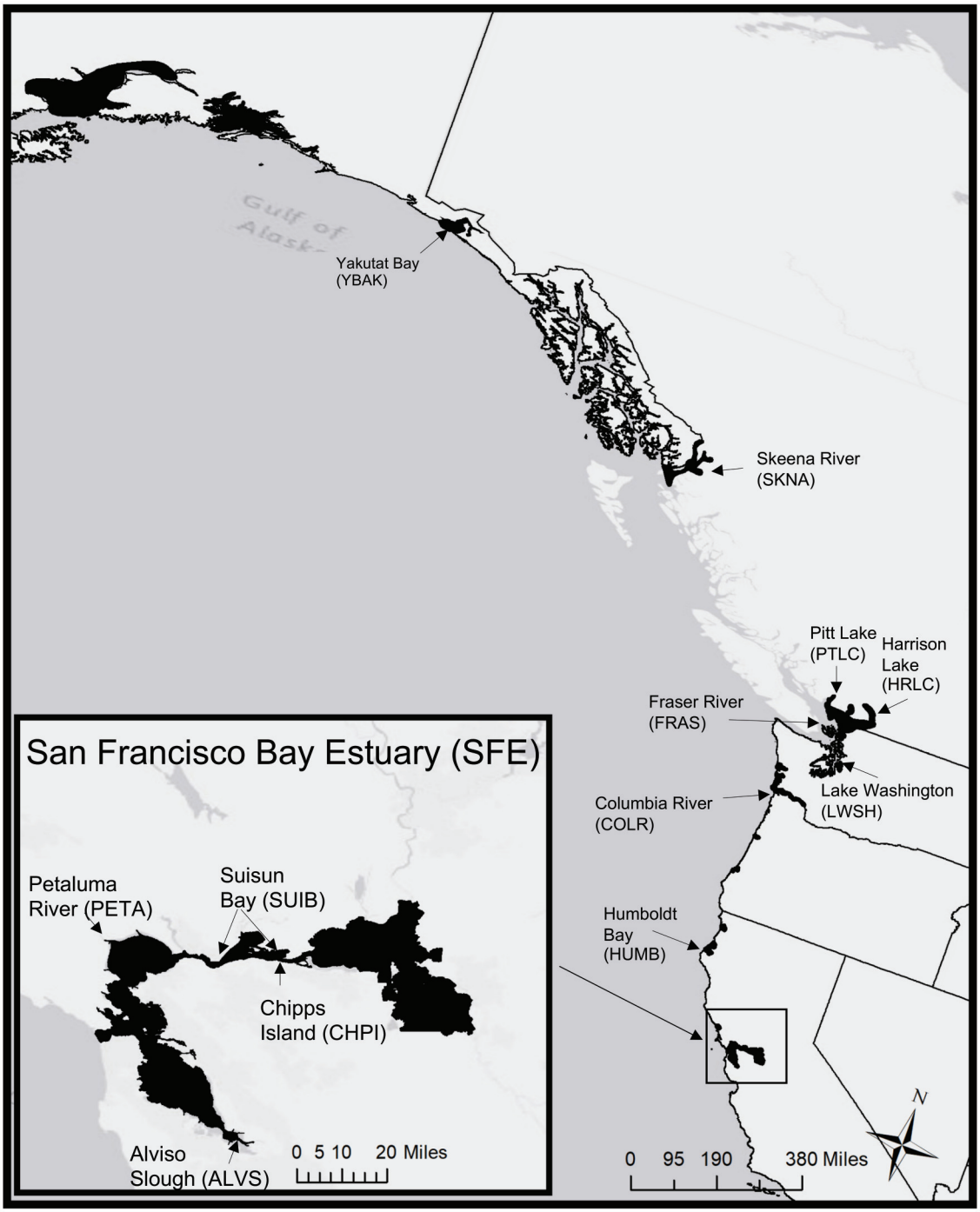
Small pelagic forage fishes form an energy conduit from primary producers to top predators and are a critical link in coastal food webs (Lenfest Forage Fish Task Force 2012; Pikitch et al. 2014). While genetic structure and connectivity among populations of commercially important forage fish species such as herring, anchovies, and sardines are relatively well-studied (Hedgecock et al. 1989), other forage fish species are poorly understood despite pressing conservation needs. Various Osmeridae "true smelt" species inhabiting the west coast of North America have shown coast-wide population declines. Eulachon (*Thaleichthys pacificus*) was once abundant in many estuaries from San Francisco Bay to the southeastern Bering Sea, but for the past two decades has declined in the southern portion of its range. The Southern Distinct Population Segment (DPS) of eulachon is listed as threatened under the Federal Endangered Species Act (ESA) (Gustafson et al. 2010, 2012; Sutherland et al. 2021). In Puget Sound, surf smelt (*Hypomesus pretiosus*), a common, recreationally harvested forage fish has precipitously declined (Greene et al. 2015), and in the San Francisco Estuary (SFE), the delta smelt (*Hypomesus transpacificus*) and the longfin smelt (*Spirinchus thaleichthys*) are possibly approaching extinction and extirpation, respectively, in the near future (Hobbs et al. 2017).

Longfin smelt is a small (fork length $<$ 150 mm) pelagic osmerid forage fish that inhabits lakes, estuaries, and marine habitats from central California to Prince William Sound, Alaska (Dryfoos 1965; Moyle 2002). The southernmost, and one of the largest, known populations occurs in the SFE (Fig. 1). Much of what is known about longfin smelt comes primarily from monitoring surveys in the San

Received 5 January 2021. Accepted 10 May 2021.

İ.K. Sağlam. Koç University, Department of Molecular Biology and Genetics, Istanbul, Turkey; Department of Animal Science, University of California, Davis, CA 95616, USA.
J. Hobbs and L.S. Lewis. Department of Wildlife, Fish and Conservation Biology, University of California, Davis, CA 95616, USA.
R. Baxter. California Department of Fish and Wildlife, Region 3, 2109 Arch-Airport Rd., Suite 100, Stockton CA 95206, USA.
A. Benjamin and A.J. Finger. Department of Animal Science, University of California, Davis, CA 95616, USA.
Corresponding authors: İsmail K. Sağlam (email: iksaglam@ucdavis.edu) and Amanda J. Finger (email: ajfinger@ucdavis.edu).
© 2021 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Fig. 1. The range-wide distribution of longfin smelt extends from central Alaska to the San Francisco Estuary, California. Documented populations depicted using filled polygons (black). The map was created using ArcMap 10.6 (Esri, Redlands, California), using base map from Esri World Ocean Base. Data for documented populations was obtained from the California Department of Fish and Wildlife's Biogeographic Information and Observation System at https://wildlife.ca.gov/Data/BIOS.



Francisco Estuary, California (SFE; Rosenfield and Baxter 2007; Baxter et al. 2008; Maunder et al. 2015; Nobriga and Rosenfield 2016; Grimaldo et al. 2017, 2020; Lewis et al. 2020) and research studies on the land-locked population in Lake Washington, King County, Washington (Moulton 1974; Sibley and Chigbu 1994; Chigbu and Sibley 1998*a*, 1998*b*; Chigbu 2000). Longfin smelt in the SFE are semi-anadromous, spawning and rearing in fresh to low-salinity waters in winter-spring, migrating to higher-salinity brackish and marine waters in early summer. Little is known about spawning behavior, though it appears that they spawn opportunistically in locations with favorable habitat conditions, with the males arriving first, followed by the females (Moyle 2002). Larvae and juvenile longfin smelt are sometimes observed in smaller estuaries indicating that smaller estuaries are utilized as spawning and rearing habitat (Lewis et al. 2020). In summer, juveniles move into higher salinity coastal waters of the Pacific Ocean, including San Francisco Bay, where they are observed as fisheries by-catch or in monitoring surveys (Rosenfield and Baxter 2007; Garwood 2017).

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Similarly, longfin smelt have been observed in the nearshore coastal environment off the Columbia River plume (Bottom and Jones 1990; Litz et al. 2014), in Bellingham Bay, Washington, near the Nooksack River (Donnelly et al. 1988; Penttila 2007), in Yakutat Bay in the discharge of two glaciers (Arimitsu 2016), and the Cook Inlet, Alaska (Abookire et al. 2000; Abookire and Piatt 2005; Jones et al. 2005). Thus, some individuals clearly spend some portion of their life history in the ocean. Documented estuarine rearing (Rosenfield and Baxter 2007) and limited detections in nearshore forage fish surveys (Penttila 2007) suggest that most fish inhabiting the coast likely remain near their natal estuary and return after a season or two on the coast (e.g., Rosenfield and Baxter 2007). Given their typical 2-year life cycle, it is unclear how long they reside in the ocean, and this may vary among populations, estuaries, or among individuals within an estuary. Little is known about the ocean-phase of their life history, and how ocean rearing and the potential for transport or movement may influence their contemporary population structure, demography, or evolutionary dynamics.

Longfin smelt was once one of the most abundant pelagic forage fish in the SFE, even contributing to a commercial bait fishery (Skinner 1962; Moyle 2002). However, based on numerous long-term monitoring surveys, the SFE population has undergone several significant declines in abundance since the 1960s (Thomson et al. 2010; Maunder et al. 2015; Nobriga and Rosenfield 2016; Hobbs et al. 2017). Longfin smelt is one of four species that declined in the early 2000s, an event known as the pelagic organism decline (POD) within the SFE (Sommer et al. 2007). The POD was attributed to anthropogenic pressures including habitat alteration, increased water diversions, food-web alteration, and invasive species (Sommer et al. 2007; Mac Nally et al. 2010; Thomson et al. 2010). In response to this decline, in 2009 the SFE population of longfin smelt was listed as threatened under the California Endangered Species Act (California Fish and Game Commission 2009), and a petition to list it as a Distinct Population Segment under the federal ESA was ruled warranted but precluded (USFWS 2012). In Oregon, longfin smelt are managed under Oregon state rule (OAR 635-004-0540 or OAR 635-994-0545) and targeted harvest off the coast is prohibited. North of Oregon there are no other protections or knowledge of long-term population status. Information about genetic diversity and structure is vital for understanding the coast-wide interactions among putative populations and for making decisions about the legal protection and conservation status of the longfin smelt in the SFE.

Here, we use genome-wide data to evaluate the population structure of longfin smelt on two scales: a broad latitudinal scale across 6 estuaries and Lake Washington, and fine scale within the Fraser River Estuary and within the SFE. We also examine genetic diversity, adaptive variation, and population connectivity in the major longfin smelt estuaries along the northeastern Pacific Coast. Our results capture patterns of local adaptation and gene flow between longfin smelt populations and highlight the need for further studies aiming to understand the population dynamics of longfin smelt.

Materials and methods

Sampling, sequencing and RAD loci discovery

We used DNA from individual longfin smelt collected between 2000 and 2015 from 11 locations along the west coast of North America (Table 1 and Fig. 1). Four of these collections were from widely dispersed locations within the SFE from long-term or targeted surveys: ALVS (Alviso Slough), CHPI (Chipps Island), PETA (Petaluma River), and SUIB (Suisun Bay). Six collections are from estuary locations further north: HUMB (Humboldt Bay, California) and COLR (Columbia River estuary, Oregon–Washington), PTLC and HRLC (Pitt Lake and Harrison Lake which are tributary lakes in the Fraser River, British Columbia), SKNA (Skeena River estuary), and YBAK (Yakutat Bay, Alaska). The last collection is from the putatively landlocked lake, Lake Washington, Washington (LWSH). Samples consisted of an individual fin clip stored in either EtOH or dried in coin envelopes. DNA was extracted using the Qiagen (Hilden, Germany) DNEasy Kit according to the manufacturer's protocol.

To produce a high-quality genomic resource for longfin smelt we sampled the genome using RAD sequencing (Miller et al. 2007; Baird et al. 2008). Paired-end 150 bp sequence reads were generated for 170 individuals using the Sfb1 restriction enzyme and the RAD protocol described in Ali et al. (2016). Sequences were sorted into individuals as matching forward and reverse fastq files using unique 8 bp barcodes. RAD-loci discovery and extension was carried out using the custom procedure given in (Sağlam et al. 2016) (for details see Supplementary information[1]). This procedure resulted in 23 525 unique RAD-contigs ranging from 300 to 800 bp and served as a de-novo reference for all downstream analyses.

Alignment and filtering

Sequence reads from individuals were aligned to the de novo set of reference RAD contigs using the BWA-MEM algorithm (Li 2013). Outputted SAM files were transformed into BAMfiles and indexed after sorting for proper pairs and removing PCR duplicates in samtools (Li et al. 2009). Potential paralogous RAD-loci were tagged using ngsParalogs (https://github.com/tplinderoth/ngsParalog) and removed from further analysis. Furthermore, we removed 47 individuals with low read depth that resulted in fewer than 100 000 alignments (Supplementary information, Table S1[1]). This left us with a final sample size of 123 individuals from 11 locations distributed among six estuaries and Lake Washington (Table 1), which formed the basis of all subsequent analyses.

SNP discovery and genotyping

Due to low sequencing depth, we conducted all analyses using the probabilistic framework implemented in ANGSD (Korneliussen et al. 2014) and related software (http://www.popgen.dk/software/index.php/Main_Page) which do not require calling genotypes and are ideal for working with low depth data (Korneliussen et al. 2013; Fumagalli et al. 2013). Genotype likelihoods, minor allele frequencies (MAF) and genotype probabilities were calculated in ANGSD. We discovered SNPs by screening for sites with an MAF of over 0.05 and the probability of the site not being polymorphic of less than 10^{-12}. Sites were discarded if they did not meet a minimum quality score of 20 and mapping quality of 10. Furthermore, we filtered out any site that had an average per individual read depth below 6 and was not present in at least half of the individuals.

Genetic structure and differentiation

We visualized genetic structure between collection sites using principal component analysis (PCA). We first estimated the genetic covariance matrix between individuals using PCAngsd based on the calculated genotype likelihoods (Meisner and Albrechtsen 2018). Principal component axes summarizing genetic structure were then derived from this matrix using classic eigenvalue decomposition in R version 3.6.1 (R Core Team 2013).

To determine shared ancestry between collection sites, we conducted admixture analyses for different values of *K*. We used *K* values ranging from 2–11. Individual admixture proportions were estimated in NgsAdmix (Skotte et al. 2013) based on the genotype likelihoods calculated in ANGSD. We ran NgsAdmix 15 times per *K* value, and used the likelihood values from each run to calculate the most likely *K* with the method described in Evanno et al. (2005).

[1]Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2021-0005.

Table 1. List of locations where longfin smelt samples were collected, including location code, number (*N*) with sequence data that passed quality filtering, year(s) collected, and approximate GPS coordinates.

Estuary	Location	*N* sampled	*N* used in analyses	Year(s) collected	GPS coordinates
San Francisco Estuary (SFBY)	Alviso Slough (ALVS), California	12	12	2013	37.4435508, –122.0055163
	Chipps Island (CHPI), California	19	7	2007, 2008, 2009	38.0551975, –121.911904
	Petaluma River (PETA), California	6	6	2015	38.1104751, –122.4874776
	Suisun Bay (SUIB), California	20	20	2008, 2010	38.07, –122.07
Humboldt Bay (HUMB)	HFAC Weir, Freshwater Creek, California	8	8	2009, 2012, 2015	40.759830294, –124.217332464
Columbia River Estuary (COLR), Oregon	Trestle Bay, North Channel, Oregon	22	8	2011	46.213389, –123.978778
Lake Washington (LWSH)	Lake Washington, Washington	20	16	2008	47.6215, –122.2558
Fraser River (FRAS)	Pitt Lake (PTLC), British Columbia	20	10	2000	49.4399889, –122.5400974
	Harrison Lake (HRLC), British Columbia	13	6	2006	49.500000, –121.833336
Skeena (SKNA)	Skeena River, British Columbia	10	10	2011	54.1499994, –129.9499962
Yakutat Bay, Alaska (YBAK)	Yakutat Bay, Alaska	20	20	2013	59.725663764, –139.836496654
Total		170	123		

Genetic differentiation between all collection sites was summarized using pairwise F_{ST} statistics. Genome-wide F_{ST} values were estimated in ANGSD by first calculating the joint-SFS between each collection site and then calculating the global weighted F_{ST} using the REALSFS module.

Genetic diversity

Genetic diversity was determined by calculating average pairwise nucleotide differences (θ_π) based on the global site frequency spectrum (SFS) derived separately for each collection location (Supplementary information, Fig. S1[1]). We used a folded SFS since we did not have a suitable outgroup for determining ancestral states. An SFS for each population was estimated in ANGSD (doSaf 1) and later used as a prior to calculate θ_π for each RAD-contig (doThetas). θ_π values for each RAD-contig were transformed into per-site values by dividing each statistic by the length of the contig and genome-wide values were obtained for each population by averaging across RAD-contigs.

Contemporary migration

We tested ongoing migration between coastal collection sites with samples grouped according to genetic structure analyses (SKNA, PTLC, HRLC, COLR, HUMB, and SFBY) using BayesAss version 3 (Wilson and Rannala 2003), which uses multilocus genotypes to determine immigrant ancestry of individuals within the last three generations (for longfin smelt, generation time is ~2 years). We called genotypes from polymorphic sites using genotype probabilities previously estimated in ANGSD using a posterior cut-off of 95% along with the major and minor allele at each site. To minimize the effect of missing genotypes and linkage we filtered out any SNP that had more than 10% missing data (1 or 2 non-called genotypes per collection location) and then chose only a single SNP per RAD contig. This resulted in a final data set of 872 SNPs which were used in migration rate estimates. We performed three independent runs with 40 million MCMC iterations each and a sampling frequency of 4000. After initial trial runs we settled at delta values 0.90, 0.90, 0.90 for parameters 1, 3 and 4 (migration rates, allele frequencies and inbreeding coefficients) which resulted in adequate mixing (40%–50%). The first 4 000 000 iterations of each run were discarded as burn-in and after checking convergence of parameters using TRACER version 1.6 (Rambaut et al. 2018) we combined all independent runs for final parameter estimation.

Local selection scans

To examine adaptive variation among sample locations, we conducted a genome-wide selection scan as implemented in PCAngsd (selection 1). Based on the FastPCA model (Galinsky et al. 2016), PCAngsd performs a selection scan along all significant PCs and identifies variants whose differentiation along top PCs is significantly greater than the null distribution of genetic drift. The advantage of this method is that it can account for population structure and gene flow between subpopulations and does not require a specific model for detecting signals of genome-wide significance (Galinsky et al. 2016). The resulting selection statistic is χ^2-distributed with 1 degree of freedom and is used to identify variants under selection based on an appropriate probability cutoff value. To minimize any confounding effects due to inbreeding we filtered out SNPs that showed significant deviations from HWE due to inbreeding (inbreedSites) based on the kinship matrix between all individuals (inbreed 3) as calculated in PCAngsd. This gave us a final set of 13 296 SNPs which resulted in a genome-wide significance threshold of 9.40×10^{-7} ($\alpha = 0.05$; 13 296 SNPS × 4 principal components).

Results

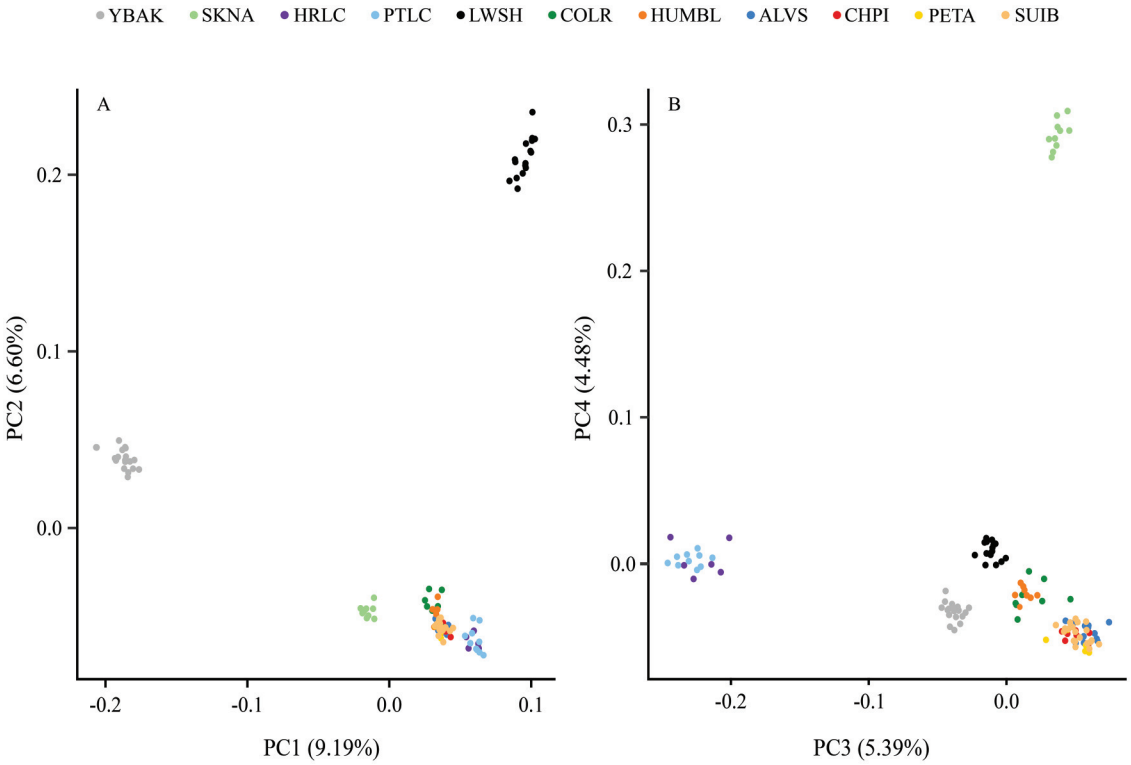
Alignment of raw reads to the de novo reference resulted in an average mapping success of 65%, and mean individual read depth was around ~5×–6×. Raw read counts, number of raw alignments, and number of alignments after filtering for proper pairs and removing PCR duplicates are given in supplementary information, Supplementary Table S1[1]. Across all collection sites, we discovered 5864 polymorphic loci and 13 402 high probability ($P < 10^{-12}$) SNPs with a minor allele frequency over 0.05.

Genetic structure and differentiation

Genetic structure for longfin smelt throughout their range was high as sampling locations along the northeastern Pacific coast were clearly separated from one another along the four significant PC axes (Fig. 2). PC1 indicated that YBAK is the most differentiated (Fig. 2), followed by LWSH on PC2, and the Fraser River sites, PTLC and HRLC, on PC3. PC3 also separated COLR and HUMB from the four SFE locations. However, we did not capture any separations among the four SFE locations (ALVS, CHPI, PETA and SUIB), between the Fraser locations, nor between COLR and HUMB along any of the significant PC axes (Fig. 2). PC4 separated SKNA from all other locations.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Fig. 2. Genetic structure of longfin smelt along the northeastern Pacific coast as summarized by PCA for comparisons between the four significant principal component axes. PC1–PC2 (A), PC3–PC4 (B). Proportion of variance explained by each PC axis is given in parentheses. Sample codes: YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); ALVS (Alviso Slough); CHPI (Chipps Island); PETA (Petaluma River); and SUIB (Suisun Bay). [Colour online.]



Admixture analysis returned support for groupings of K = 3 and K = 5 based on Evanno's method (Supplementary information, Fig. S2[1]). Admixture plots based on K = 3 gave a similar structure to that defined along PC1 and PC2 separating both YBAK and the landlocked LWSH while indicating some amount of shared ancestry between all other coastal sites (Fig. 3). At K = 5 the analysis was able to separate both the Skeena River (SKNA) and Fraser River locations (PTLC and HRLC) from COLR, HUMB, and the four SFE locations (Fig. 3). Similar to PCA we were not able to capture any further structure between sites at higher K values, and could not distinguish between the four SFE locations, the Fraser locations, or COLR and HUMB (Supplementary information, Fig. S3[1]).

Genetic differentiation based on genome-wide F_{ST} values supported results obtained from structure analysis. Relatively high genetic differentiation ($F_{ST} > 0.1$) was observed for pairwise comparisons featuring YBAK, SKNA and LWSH, while pairwise F_{ST} values between PTLC and HRLC and other locations were moderate ($0.09 > F_{ST} > 0.1$), and other comparisons were generally low ($0.021 > F_{ST} > 0.052$; Fig. 4).

Due to low differentiation and multiple sampling sites in the Fraser River and the SFE, we conducted separate PCA and admixture analysis on just these samples to test for the presences of fine scale population structure. Separate PCA and admixture analysis did not reveal any fine scale structure between SFE locations (Supplementary information, Fig. S4[1]); however, we were able to capture significant structure within the Fraser River (Supplementary information, Fig. S5[1]). Based on genetic structure and differentiation we conclude that longfin smelt within SFE constitute one single population (hereinafter called SFBY).

Contemporary migration

Migration rates between locations (except LWSH) based on all three BayesAss runs were consistent with each run and converged onto the same parameter space (Supplementary information, Fig. S6[1]), so we combined all three runs for final parameter estimation using approximately 30 000 posterior samples. With two exceptions, we observed relatively low migration probabilities between populations (<0.025), with the lower 95% bayesian credible intervals (highest probability densities, HPD) nearing zero, indicating limited migration between most locations (Fig. 5; Supplementary information, Table S2[1]). However, we did find significant evidence for contemporary migration northward from SFBY to both HUMB and COLR with probabilities of 0.199 (HPD: 0.119–0.277) and 0.201 (HPD: 0.118–0.275) migrants per generation respectively (Table S2[1]). However, migration probabilities in the reverse direction (south from HUMB and COLR to SFBY) were low and not significant. Similarly, we found significant migration from PTLC east and upstream to HRLC (0.191, HPD: 0.109–0.273) but none in the reverse direction (Fig. 5; Supplementary Table S2[1]).

Genetic diversity and local selection scans

Average θ_π was similar between all populations and ranged from 0.001 to 0.003 (Fig. 6). HRLC had the highest value, followed by PTLC, COLR, and SFBY. To find evidence of adaptive variation within populations we conducted selection scans to find loci that show significant differentiation between populations above that expected by genetic drift and accounted for by structure. Selection scans conducted on the four significant PC axes did not capture any sites that reached genome-wide-significance along PC1 or PC2, but captured 21 and 23 sites along PC3 and PC4, respectively (Supplementary information, Tables S3 and S4[1]). To understand changes in allele frequencies at these sites we plotted the frequency of the reference allele (taken as the minor allele in the entire sample) at these sites for all populations. Along PC3 (which separates the Fraser River locations PTLC and HRLC), the frequency of the reference allele at sites with

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Fig. 3. Shared ancestry between longfin smelt along the northeastern Pacific coast as revealed by admixture analysis for $K = 3$ and $K = 5$. YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); ALVS (Alviso Slough); CHPI (Chipps Island); PETA (Petaluma River); and SUIB (Suisun Bay). [Colour online.]

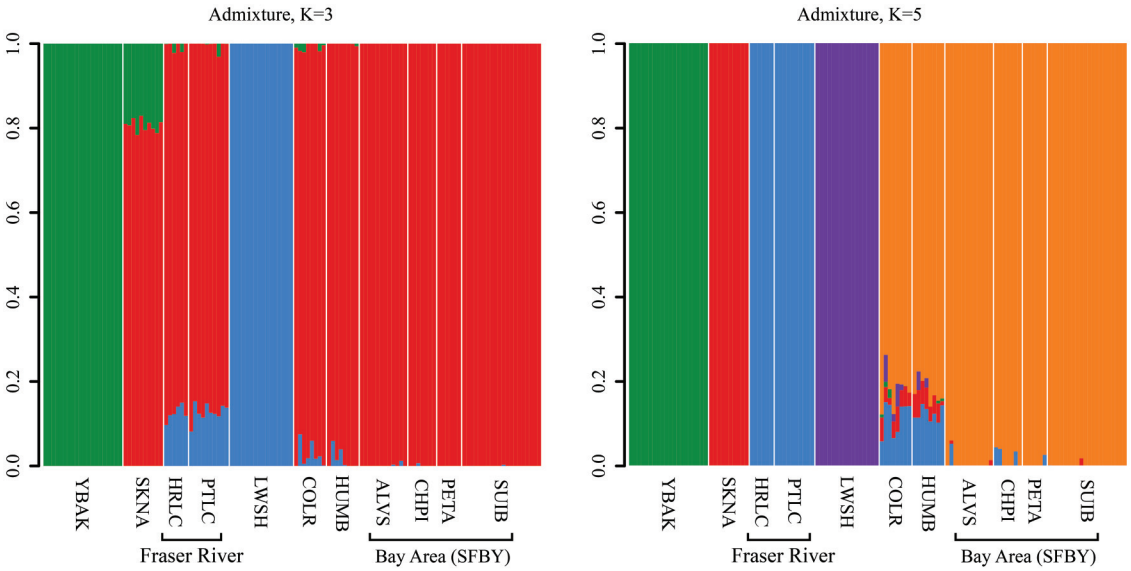
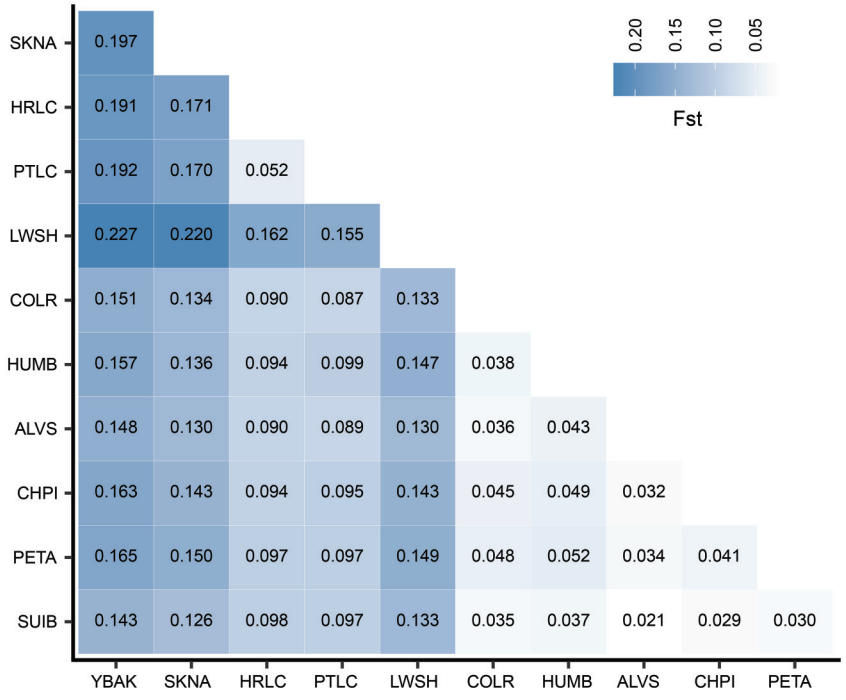


Fig. 4. Genetic differentiation between longfin smelt collections along the northeastern Pacific coast as measured by F_{ST} statistics. YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); ALVS (Alviso Slough); CHPI (Chipps Island); PETA (Petaluma River); and SUIB (Suisun Bay). [Colour online.]



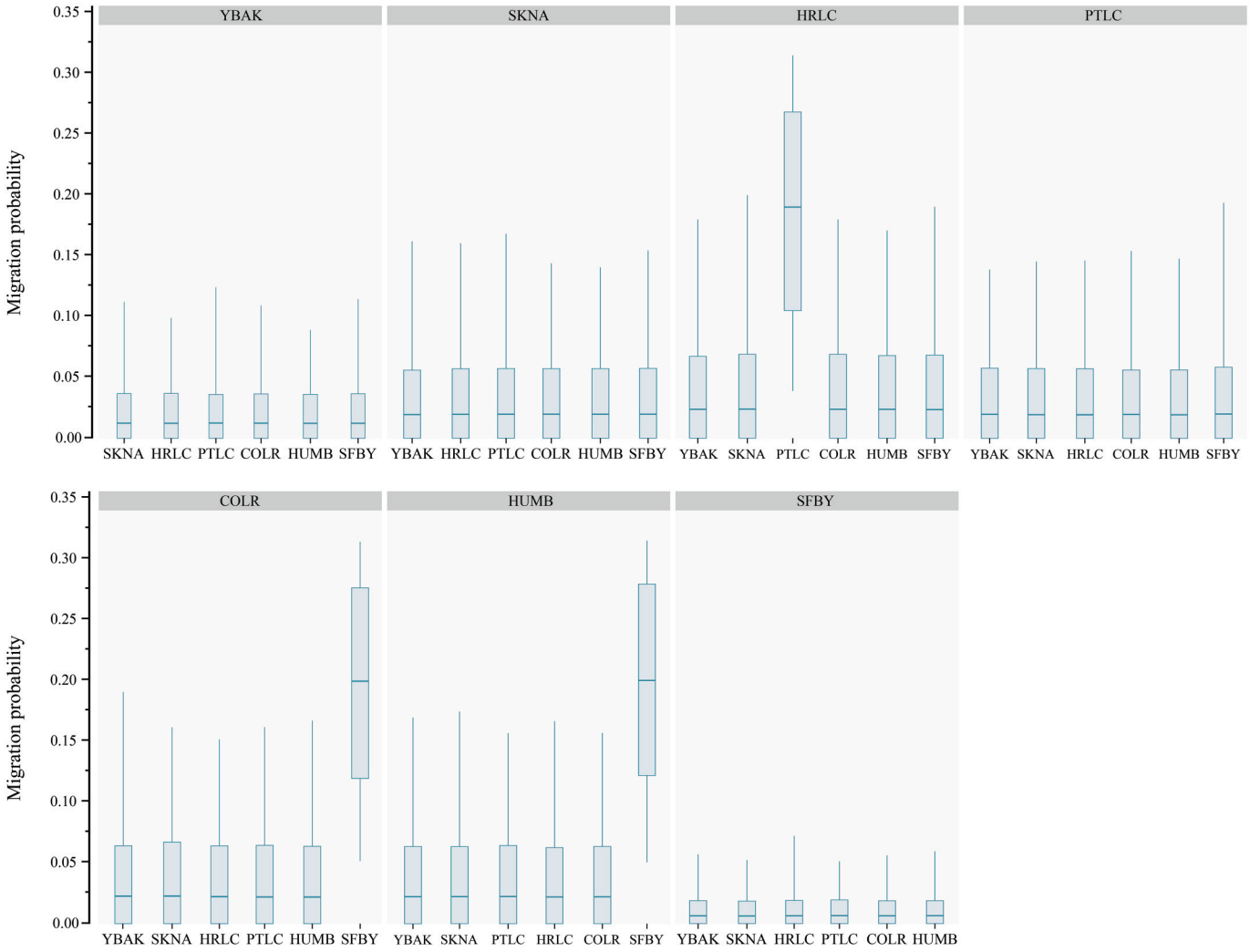
positive signatures of selection were higher in the Fraser locations, and some sites were swept to near fixation (Fig. 7A). Likewise, along PC4 (which separates SKNA), the frequency of the reference allele at sites showing signs of selection were higher in the SKNA population, to the point where the minor allele becomes the major allele (Fig. 7B).

Discussion

This is the first range-wide genomic analysis of the genetic structure within longfin smelt, an understudied forage fish. Our results indicate relatively high genetic structure between the major estuaries sampled, fine-scale structure between Pitt and Harrison lakes in

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Fig. 5. Depiction of ontemporary migration rates (calculated as 95% Bayesian credible intervals in the software program BayesAss) between collections of longfin smelt, as probability of migrants received within the last three generations in each population. Each box represents a different sink population, and the code locations below indicate sources. YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); SFBY (San Francisco Estuary, includes Alviso Slough, Chipps Island, Petaluma River, and Suisun Bay). [Colour online.]



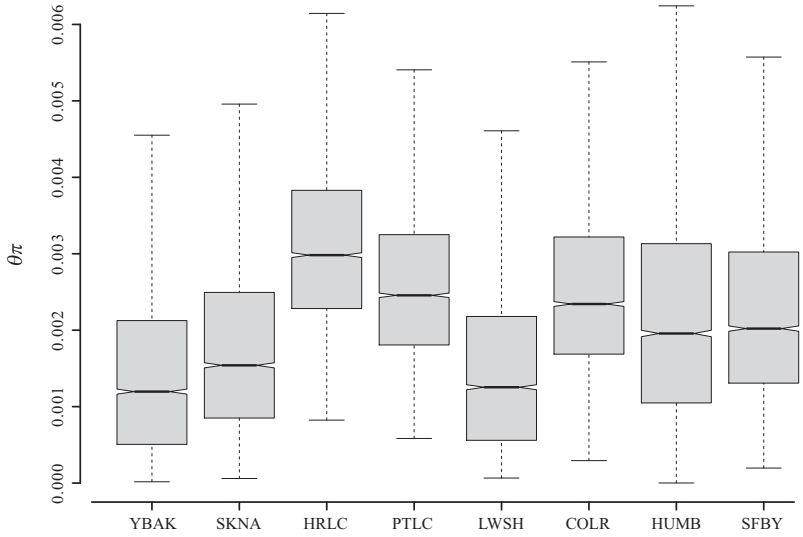
the Fraser River Estuary, and very low levels of structure within the San Francisco Estuary, by far the largest estuary in the region. Genetic structure was more pronounced between northern estuaries whereas we detected appreciable amounts of shared ancestry and ongoing gene flow in southern estuaries. Migration analyses showed unidirectional migration from Pitt to Harrison Lake in the Fraser River system and strong unidirectional northward migration out of the SFE to nearby estuaries of Humboldt Bay and Columbia River. Furthermore, we detected clear signatures of local adaptation samples taken from the Fraser River system and the Skeena River Estuary. Taken together, our results identify broad patterns of genetic diversity in longfin smelt along the North American west coast shaped by co-ancestry, unidirectional migration and local adaptation and hint at variability in longfin smelt behavior during the ocean phase of their lives.

Genetic structure, gene flow, and local adaptation in northern estuaries

Overall, the observed patterns of differentiation and low levels of gene flow among northern estuaries indicate low levels of movement between estuaries and (or) a lack of regular long-distance migration, suggesting that genetic drift had a strong influence (e.g., bottlenecks and founder events). The northernmost sampled population in Yakutat Bay was the most differentiated with the lowest genetic diversity, followed by the landlocked Lake Washington, Skeena River, and the Fraser River locations Pitt and Harrison lakes. Yakutat Bay is a relatively small bay (29 km wide) located at the southwest of the Disenchantment Bay waters which captures freshwater runoff from three glaciers, including the Hubbard Glacier. Some geomorphological patterns influencing Yakutat Bay include the advance and retreat cycles of Hubbard Glacier that create Russell Lake via temporary ice dams that then burst sending cold freshwater into Yakutat Bay, as well as high seismic activity (Barclay et al. 2001). This combined with the relatively recent Pleistocene retreat of glaciers opening up the Alaska coast may contribute to the high differentiation of the Yakutat Bay population. Interestingly, eulachon, whose range overlaps with longfin smelt have a contrasting pattern of genetic structuring; they have weak structure (Candy et al. 2015). This may be the result of the eulachons' much higher marine dependance than longfin smelt; after spawning,

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Fig. 6. Genome wide values of per site θ_π (average pairwise nucleotide differences) for longfin populations along the northeastern Pacific coast. YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); SFBY (San Francisco Estuary, includes Alviso Slough, Chipps Island, Petaluma River, and Suisun Bay).



Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

eulachon larvae are immediately flushed out to sea and dispersed passively by current (Gustafson et al. 2012; Hay and McCarter 2000).

The provenance of longfin smelt in Lake Washington is somewhat mysterious. Lake Washington is a shallow (~74 m) ribbon lake formed by glaciation processes. The original outlet was the Black River, which evidently had a salmon run and was somewhat navigable to Puget sound depending on tides and conditions (Purvis 1934, Chrzastowski 1983). In 1916 a ship canal was built forming the current outlet and causing the lake level to drop 3 m and dry the Black River. The ship canal possesses locks with fish ladders, and the canal is successfully used by salmon to access Lake Washington, but it is unknown whether longfin smelt can use either the ladders or navigate the locks. The Lake Washington population is highly differentiated; it is possible that it is remnant from before the locks were built, or perhaps they were able to colonize Lake Washington later.

Though long-distance movement of longfin smelt between estuaries seems to be broadly limited in northern estuaries, migration analyses indicated short-distance, unidirectional gene flow from Pitt Lake to Harrison Lake. The findings from Pitt and Harrison lakes are supported by their hydrology as they are fjord lakes that connect to the Fraser River. Harrison Lake is 60 km to the east and upstream of Pitt Lake and has a greater connection to the Fraser River due to higher average flows (446 $m^3 \cdot s^{-1}$) draining a substantially larger watershed (8324 km^2) compared to the Pitt watershed (1660 km^2 and 186 $m^3 \cdot s^{-1}$). In addition, Harrison River was dredged to make it navigable to boats entering Harrison Lake. In contrast, Pitt Lake is one of the largest tidal fjord lakes in the world. Very strong tides create a negative delta that is exposed at low tide, cutting it off from the Pitt River which empties into the Fraser River during part of each day (Ages and Woollard 1976; Fisheries and Oceans Canada Habitat and Enhancement Branch 1999). Interestingly, reports on longfin smelt have suggested that Pitt and Harrison lakes are land-locked (Dryfoos 1965, USFWS 2012), though our migration analysis suggests that they are not.

There is evidence that local adaptation might have played a role in the Skeena River and Pitt and Harrison lakes. Examination of local adaptation using genome-wide selection scans along significant principal component axes did not reveal any regions reaching genome-wide significance along the two main axes (PC1 and PC2) suggesting that differentiation between estuaries is mainly caused by neutral evolution. However, selection scans along PC3 and PC4 revealed a large number of sites that reached genome wide significance in Pitt and Harrison lakes (N = 21 sites along PC3) and the Skeena River (N = 23 sites along PC4) indicating local adaptation. In all three systems the frequency of the outlier allele had risen to over 60% (going to total fixation in some) indicating strong selective pressures on these genomic regions. The colder and more oligotrophic freshwater habitats of Pitt and Harrison lakes in the Fraser River drainage may provide very different challenges than the mostly anadromous life histories of other locations (Dryfoos 1965) and could be the source of selection in this system. Interestingly, Dryfoos (1965) described Harrison Lake longfin smelt as "stunted" and subsequent descriptions have described populations in both Pitt and Harrison lakes as "pygmy" longfin smelt, although Dryfoos (1965) attributed their small size to the colder low-food environment. Future studies may determine if this phenotype has a genetic basis. In the Skeena River estuary, the limited knowledge of longfin smelt presents no obvious sources of strong selection, though the estuary is one of the more complex, with multiple channels, and several islands that both dissect the estuary and enclose it.

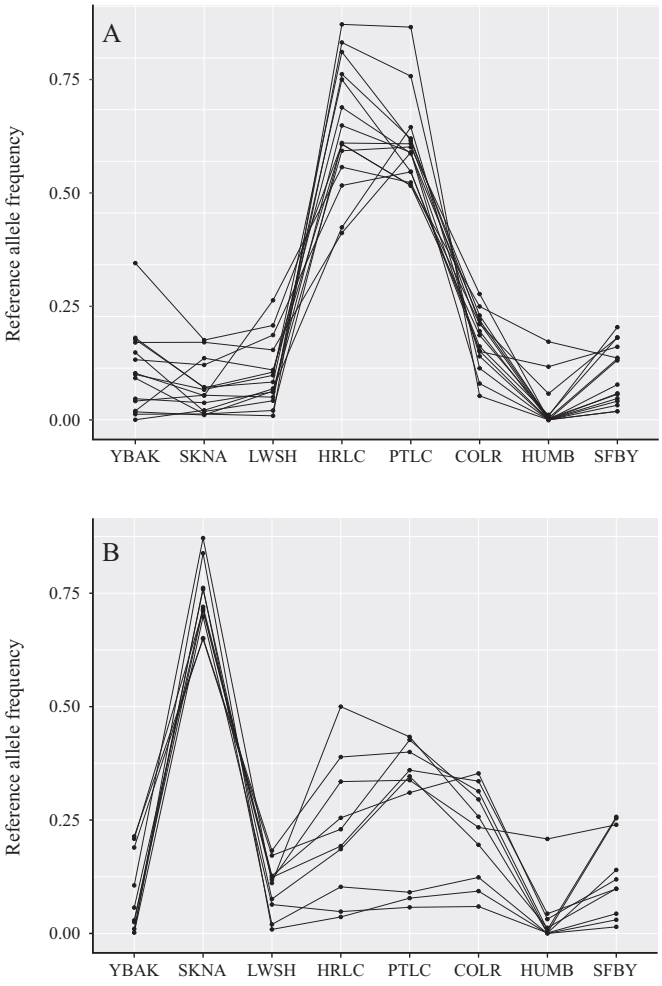
Further studies are needed to confirm patterns of local adaptation especially if such patterns exist in other estuaries as well. Since our sample sizes were fairly small and we interrogated a very small portion of the genome (12 695 loci) there is a good chance that we were only able to capture loci under extreme selection pressures linked to traits with high heritability. In addition, this method of selection scanning cannot detect all types of adaptive variation, such as rare variants, and cannot indicate how or when selection occurred, only that it occurred. Therefore, absence of local adaptation in other estuaries might result from limitations of the current study and not necessarily the absence of selection pressures. Greater sampling coverage may uncover additional adaptive variation and present further avenues for research.

Reduced genetic structure and northward migration in southern estuaries

In contrast to northern estuaries, the locations sampled at the southern end of our sampling range showed reduced genetic structure and differentiation, and appreciable amounts of shared ancestry, especially between the SFE, Humboldt Bay and Columbia River estuaries. Though this pattern may be due to recent divergence, it is also supported by the detection of relatively high levels of

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Fig. 7. Reference allele frequencies of sites under selection along PC3 (A) and PC4 (B). YBAK (Yakutat Bay, Alaska); SKNA (Skeena River estuary); HRLC (Harrison Lake, Fraser River, British Columbia); PTLC (Pitt Lake, Fraser River, British Columbia); LWSH (Lake Washington); COLR (Columbia River estuary, Oregon–Washington); HUMB (Humboldt Bay, California); SFBY (San Francisco Estuary, includes Alviso Slough, Chipps Island, Petaluma River, and Suisun Bay).



ongoing northward gene flow from the SFE towards Humboldt Bay and the Columbia River estuary.

Evidence of contemporary gene flow northward out of the SFE raises a number of questions, the first of which is: what is the process by which genetic influence is moving north? Our findings support contemporary gene flow out of the SFE to estuaries that are a considerable distance north: Humboldt Bay is 418 sea km and the Columbia River is 1030 sea km north of the SFE. There are several smaller estuaries north of the SFE where longfin smelt have been found, including the Russian, Eel, and Klamath rivers in California (Garwood 2017), and in Coos Bay, Tillamook Bay, and the Yaquina Bay estuaries in Oregon (Bottom and Forsberg 1978); it is possible that longfin "estuary hop" in a northern direction rather than undergo long distance migrations. In the Columbia River Estuary, an understanding of the current distribution and abundance trends of longfin smelt is limited. Humboldt Bay is a coastal lagoon, and the second largest estuary on the California coast. Due to low freshwater inputs, Humboldt Bay itself likely does not support (and may have never supported) a large population, but in recent years small collections have occurred regularly in bay tributaries (Garwood 2017). Alternatively, a local population in the larger Eel River estuary located just south may support the common occurrence of longfin smelt in Humboldt Bay (Pequegnat and Butler 1982; Garwood 2017). Nevertheless, longfin smelt were considered common in Humboldt Bay in the late 1970s, though the population has likely declined at least since then, if not over the last 150 years when habitat alteration began (Garwood 2017; USFWS 1994). Given its small size and human alteration, it is not surprising that longfin smelt captured in Humboldt Bay today would exhibit genetic influence from other locations.

The second major question regarding gene flow is: why is gene flow only detected in a northern direction from the SFE? Longfin smelt have been found offshore of the SFE during winter months (Garwood 2017), when the near-shore Davidson current is flowing northward, and providing conditions for modest northward movement annually. In extreme years associated with strong El Niños, nearshore organisms can be transported multiple hundreds of kilometers northward as seen in range extensions of demersal species with pelagic egg and larval stages (e.g., California tonguefish; Huyer and Smith 1985; Pearcy et al. 1985). In contrast, the California current which generally flows southward and far offshore during spring through fall, is highly complex, and varies in direction and strength at different depths, distances from shore, and times of the year. Thus, the California current may be less likely to assist southward gene flow. Moreover, upwelling areas in the Pacific are notoriously variable and include seasonal, annual, and decadal scale changes in oceanographic patterns (Checkley and Barth 2009). These upwelling changes may cause punctuated episodes of gene flow, and have been linked to abundance and range shift variation along the coast of California in other pelagic forage fish that commonly range well offshore: Pacific sardine (*Sardinops sadax caerulea*) and northern anchovy (*Engraulis mordax*) along the coast of California (Lecomte et al. 2004). We cannot determine if there is gene flow due to spawning in smaller estuaries in between the estuaries we sampled, or if longfin smelt actually migrate long distances, or what scale gene flow occurs on.

Lack of fine-scale structure in the San Francisco Estuary

Due to the large size and complexity of the SFE, it is reasonable to expect fine scale genetic structuring of longfin smelt taken from different sampling locations within the SFE, especially if fish exhibit fidelity to unique spawning and rearing habitats. However, our analysis shows minimal structure, strongly suggesting that the whole estuary is a single genetic unit. Though long-term data sets do not directly target longfin smelt, they have been leveraged by several studies to understand the ecology of longfin smelt in the SFE. For example, several studies link abundance to various conditions in the estuary, such as freshwater flows (Jassby et al. 1995; Kimmerer 2002; Grimaldo et al. 2020). Rosenfield and Baxter (2007) synthesized three long term data sets to investigate patterns in longfin smelt abundance and presence throughout their life cycle. Based on declines in detections (and abundance) of the second year of life to nearly zero in late summer and fall followed by increasing detections in winter, Rosenfield and Baxter (2007) concluded that most longfin smelt must spend the months (June–August) in the coastal ocean. However, given their 2- to 3-year lifespan, they can reside at sea for more than a year (Rosenfield and Baxter 2007). A couple age classes have been collected in the coastal ocean in the past (CDFG 2009) indicating the potential for coastal rearing beginning their first summer and continuing until returning to the estuary to spawn at the end of their second fall or the following winter (2-year lifespan), or an entire year later (3-year lifespan). They also suggested that anadromy may be an alternative life history strategy, with some fish staying in the estuary as residents. Such residents likely rear in cool marine waters in the lower estuary during the late-summer and fall of their second year, making salinity distinctions between residents and coastal migrants difficult. Otolith microchemistry using isotopes of strontium and oxygen further supports these observations, indicating that individual longfin

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

smelt within the SFE can exhibit a variety of spawning, rearing, and adult migratory behaviors (Lewis et al. 2020).

Management implications for SFE

Results of this work have several key implications that can help guide conservation efforts for longfin smelt in the SFE. The presence of unidirectional northward migration from the SFE to nearby estuaries further emphasizes the importance of the SFE for longfin smelt populations throughout their southern range, and simultaneously raises concern about the vulnerability of the SFE population to long-term extirpation, because the population is apparently not supported by a nearby source population. Together, these results suggest that protection of the SFE population is likely warranted.

The lack of significant population structure within the SFE was somewhat surprising given its geographic size; however, if longfin smelt do not exhibit unique spawning habitats with high site fidelity, then this result might be expected. Given that the majority of low-salinity spawning and rearing habitat appears to be located in the upper estuary ("Delta"; Moyle 2002); it is possible that nearly all longfin smelt spawn within the same geographic region and later disperse throughout the estuary as late-stage larvae and juveniles. Though longfin smelt have been detected spawning and rearing in and near tributaries located throughout the SFE, especially during years of high freshwater outflow (Lewis et al. 2020), the lack of genetic structure may further emphasize the importance of SFE's Delta as critical spawning habitat. The Delta is highly impacted by water diversions, invasive species, pesticides, and many other anthropogenic factors that may be addressed by well-informed policy decisions (Sommer et al. 2007, Baxter et al. 2010). However, before strong conclusions can be made regarding fine-scale population structure and subsequent policies drafted, additional studies are needed examining genomic variation using larger sample sizes collected over longer time scales.

Conclusion

The significant, broad-scale genetic structure and diversity of longfin smelt in the northeastern Pacific Ocean is likely shaped by a multitude of factors including strong genetic drift, local adaptation, shared ancestry, and gene flow. Given the variety of coastal and freshwater habitats used by longfin smelt throughout its range, and the reported diversity of life history types, it is unsurprising to find evidence suggesting that different dynamics govern the genetic structure of separate populations. The most interesting results of this work were the observation of genetic isolation across broad latitudinal scales and the findings of unidirectional, long-distance northward gene flow out of the SFE into other southern estuaries. Though this latter observation challenges previous assumptions of a limited migratory capacity for this species, it is in agreement with the broad coastal distribution of the species and the complex nearshore coastal oceanography of the region. Further genomic work using larger sample sizes collected over more locations and longer time scales may provide additional insights regarding the genetic structure of this population. We also suggest combining genomics with other tools such as otolith chemistry to shed light on variation in migratory behaviors.

Acknowledgements

The authors thank Colin Grant (US Fish and Wildlife Service), who made many of the initial inquiries soliciting longfin smelt samples, and Josh Israel (US Bureau of Reclamation), who performed initial genetic analyses. The authors also thank the following individuals and agencies for help collecting samples: Hobbslab and staff – UC Davis; Matt Dekar and staff – US Fish and Wildlife Service; Kathy Hieb, Jennifer Giannetta, Jeremiah Bautista, Aaron Ng, Rebecca Garwood, James Ray, Justin Garwood, and Mike Wallace – California Department of Fish and Wildlife; Laurie Lloyd, Olaf Langness – Washington Department of Fish and Wildlife, Thomas Quinn, University of Washington; Eric Taylor – The University of British Columbia, John Kelson – Fish Biologist–Consultant, Smithers, British Columbia; Mayumi Arimitsu – US Geological Survey, Alaska Science Center. Finally, the authors also thank a number of internal reviewers, Genomic Variation Lab Personnel, and other agency biologists who helped this project, including Matthew Campbell, Daphne Gille, Levi Lewis, Andrea Schreier, Josh Israel, and Alisha Goodbla. Funding for this project was generously provided by the California Department of Water Resources, Agreement Number 4600011196.

References

Abookire, A.A., and Piatt, J.F. 2005. Oceanographic conditions structure forage fishes into lipid-rich and lipid-poor communities in lower Cook Inlet, Alaska, USA. Mar. Ecol. Prog. Ser. **287**: 229–240. doi:10.3354/meps287229.

Abookire, A.A., Piatt, J.F., and Robards, M.D. 2000. Nearshore fish distributions in an Alaskan estuary in relation to stratification, temperature and salinity. Estuar. Coast. Shelf Sci. **51**(1): 45–59. doi:10.1006/ecss.1999.0615.

Ages, A., and Woollard, A. 1976. The tides in the Fraser Estuary. Pacific Marine Science Report 76-5. Institute of Ocean Sciences, Patricia Bay, Victoria, B.C. pp. 7, 11.

Ali, O.A., O'Rourke, S.M., Amish, S.J., Meek, M.H., Luikart, G., Jeffres, C., and Miller, M.R. 2016. RAD capture (rapture): flexible and efficient sequence-based genotyping. Genetics, **202**(2): 389–400. doi:10.1534/genetics.115.183665. PMID:26715661.

Arimitsu, M. 2016. The influence of glaciers on coastal marine ecosystems in the Gulf of Alaska. PhD dissertation, University of Alaska, Fairbanks, Alaska.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE, **3**(10): e3376. doi:10.1371/journal.pone.0003376. PMID:18852878.

Barclay, D.J., Calkin, P.E., and Wiles, G.C. 2001. Holocene history of Hubbard Glacier in Yakutat Bay and Russell Fiord, southern Alaska. Geol. Soc. Am. Bull. **113**(3): 388–402. doi:10.1130/0016-7606(2001)113<0388:HHOHGI>2.0.CO;2.

Baxter, R., Breuer, R., Brown, L., Chotkowski, M., Feyrer, F., Gingras, M., et al. 2008. Pelagic Organism Decline Progress Report: 2007 synthesis of results. Interagency Ecological Program for the Sacramento-San Joaquin Estuary Technical Report. Available from https://www.waterboards.ca.gov/waterrights/water_issues/programs/bay_delta/pelagic_organism/docs/pod_ieppodmt_2007synthesis_011508.pdf.

Baxter, R., Breuer, R., Brown, L., Conrad, L., Feyrer, F., Fong, S., et al. 2010. Interagency Ecological Program 2010 Pelagic Organism Decline Work Plan and Synthesis of Results. Interagency Ecological Program for the San Francisco Estuary.

Bottom, D., and Forsberg, B. 1978. The fishes of Tillamook Bay. Federal Aid Progress Reports Fisheries. Oregon Department of Fish and Wildlife, Portland, Ore.

Bottom, D.L., and Jones, K.K. 1990. Species composition, distribution, and invertebrate prey of fish assemblages in the Columbia River Estuary. Prog. Oceanogr. **25**(1–4): 243–270. doi:10.1016/0079-6611(90)90009-Q.

California Fish and Game Commission. 2009. California fish and game commission notice of findings: longfin smelt (*Spirinchus thaleichthys*). Meeting, Woodland, Calif., 29 May, 2009.

Candy, J.R., Campbell, N.R., Grinnell, M.H., Beacham, T.D., Larson, W.A., and Narum, S.R. 2015. Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. Mol. Ecol. Resourc. **15**(6): 1421–1434.

CDFG. 2009. Report to the Fish and Game Commission: A status review of the longfin smelt (*Spirinchus thaleichthys*) in California. Report to the Fish and Game Commission. California Department of Fish and Game.

Checkley, D.M., Jr., and Barth, J.A. 2009. Patterns and processes in the California Current system. Prog. Oceanogr. **83**(1–4): 49–64. doi:10.1016/j.pocean.2009.07.028.

Chigbu, P. 2000. Population biology of longfin smelt and aspects of the ecology of other major planktivorous fishes in Lake Washington. J. Freshw. Ecol. **15**(4): 543–557. doi:10.1080/02705060.2000.9663777.

Chigbu, P., and Sibley, T.H. 1998*a*. Feeding ecology of longfin smelt (*Spirinchus thaleichthys* Ayres) in Lake Washington. Fish. Res. **38**(2): 109–119. doi:10.1016/S0165-7836(98)00156-8.

Chigbu, P., and Sibley, T.H. 1998*b*. Predation by longfin smelt (*Spirinchus thaleichthys*) on the mysid *Neomysis mercedis* in Lake Washington. Freshwater Biol. **40**(2): 295–304. doi:10.1046/j.1365-2427.1998.00354.x.

Chrzastowski, M. 1983. Historical changes to the Lake Washington and route of the Lake Washington Ship Canal, King County, Washington. Open-File Report 81-1182. US Geological Survey. Available from https://pubs.er.usgs.gov/publication/ofr811182.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22

Donnelly, R.F., Miller, B.S., Stadler, J.H., Christensen, L., Larsen, K., and Dinnel, P.A. 1988. Puget Sound dredge disposal analysis (PSDDA) phase II disposal site bottomfish investigations: final report. Fisheries Research Institute, School of Fisheries, University of Washington, Seattle, Wash.

Dryfoos, R.L. 1965. The life history and ecology of the longfin smelt in Lake Washington. Ph.D. thesis, University of Washington, Seattle, Wash.

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.

Fisheries and Oceans Canada Habitat and Enhancement Branch. 1999. Lower Fraser Valley streams strategic review: Lower Fraser Valley stream review, Vol. 1. Fraser River Action Plan, Vancouver, B.C. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/240006.pdf.

Fumagalli, M., Vieira, F.G., Korneliussen, T.S., Linderoth, T., Huerta-Sánchez, E., Albrechtsen, A., and Nielsen, R. 2013. Quantifying population genetic differentiation from next-generation sequencing data. Genetics, **195**(3): 979–992. doi:10.1534/genetics.113.154740. PMID:23979584.

Galinsky, K.J., Bhatia, G., Loh, P.-R., Georgiev, S., Mukherjee, S., Patterson, N.J., and Price, A.L. 2016. Fast principal-component analysis reveals convergent evolution of ADH1B in Europe and East Asia. Am. J. Hum. Genet. **98**(3): 456–472. doi:10.1016/j.ajhg.2015.12.022. PMID:26924531.

Garwood, R.S. 2017. Historic and contemporary distribution of longfin smelt (*Spirinchus thaleichthys*) along the California coast. California Fish and Game, **103**(3): 96–117.

Greene, C., Kuehne, L., Rice, C., Fresh, K., and Penttila, D. 2015. Forty years of change in forage fish and jellyfish abundance across greater Puget Sound, Washington (USA): anthropogenic and climate associations. Mar. Ecol. Prog. Ser. **525**: 153–170. doi:10.3354/meps11251.

Grimaldo, L., Feyrer, F., Burns, J., and Maniscalco, D. 2017. Sampling uncharted waters: examining rearing habitat of larval longfin smelt (*Spirinchus thaleichthys*) in the upper San Francisco Estuary. Estuar. Coast. **40**(6): 1771–1784. doi:10.1007/s12237-017-0255-9.

Grimaldo, L., Burns, J., Miller, R.E., Kalmbach, A., Smith, A., Hassrick, J., and Brennan, C. 2020. Forage Fish Larvae Distribution and Habitat Use During Contrasting Years of Low and High Freshwater Flow in the San Francisco Estuary. San Francisco Estuary and Watershed Science, **18**(3). doi:10.15447/sfews.2020v18iss3art5.

Gustafson, R.G., Ford, M.J., Teel, D., and Drake, J.S. 2010. Status review of eulachon (*Thaleichthys pacificus*) in Washington, Oregon, and California. NOAA Tech. Memo. NMFS-NWFSC-105. US Department of Commerce, Washington, D.C.

Gustafson, R.G., Ford, M.J., Adams, P.B., Drake, J.S., Emmett, R.L., Fresh, K.L., et al. 2012. Conservation status of eulachon in the California Current. Fish. Fish. **13**(2): 121–138. doi:10.1111/j.1467-2979.2011.00418.x.

Hay, D.E., and McCarter, P.B. 2000. Status of the eulachon *Thaleichthys pacificus* in Canada. Research Document 2000-145. Department of Fisheries and Oceans Canada, Canadian Stock Assessment Secretariat, Ottawa.

Hedgecock, D., Hutchinson, E.S., Li, G., Sly, F.L., and Nelson, K. 1989. Genetic and morphometric variation in the Pacific sardine, *Sardinops sagax caerulea*: comparisons and contrasts with historical data and with variability in the northern anchovy, *Engraulis mordax*. Fish. Bull. **87**(3): 653–671.

Hobbs, J.A., Moyle, P.B., Fangue, N., and Connon, R.E. 2017. Is extinction inevitable for delta smelt and longfin smelt? An opinion and recommendations for recovery. SFEWS. **15**(2): Article 2. doi:10.15447/sfews.2017v15iss2art2.

Huyer, A., and Smith, R.L. 1985. The signature of El Nino off Oregon, 1982–1983. J. Geophys. Res. **90**(C4): 7133–7142. doi:10.1029/JC090iC04p07133.

Jassby, A.D., Kimmerer, W.J., Monismith, S.G., Armor, C., Cloern, J.E., Powell, T.M., Schubel, J.R., and Vendlinski, T.J. 1995. Isohaline position as a habitat indicator for estuarine populations. Ecol. Appl. **5**: 272–289.

Jones, T.M., Bennett, A.L., and Hamon, T.R., 2005. Baseline inventory of freshwater fishes of the Southwest Alaska Inventory and monitoring network: Alagnak Wild River, Aniakchak National Monument and Preserve, Katmai National Park and Preserve, Kenai Fjords National Park and Lake Clark National Park and Preserve. National Park Service, Anchorage, Alaska. Available from https://www.arlis.org/docs/vol1/71090809.pdf.

Kimmerer, W.J. 2002. Effects of freshwater flow on abundance of estuarine organisms: physical effects or trophic linkages? Mar. Ecol. Prog. Ser. **243**: 39–55.

Korneliussen, T.S., Moltke, I., Albrechtsen, A., and Nielsen, R. 2013. Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics, **14**: 289. doi:10.1186/1471-2105-14-289. PMID:24088262.

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics, **15**: 356. doi:10.1186/s12859-014-0356-4. PMID:25420514.

Lecomte, F., Grant, W.S., Dodson, J.J., Rodriguez-Sanchez, R., and Bowen, B.W. 2004. Living with uncertainty: genetic imprints of climate shifts in East Pacific anchovy (*Engraulis mordax*) and sardine (*Sardinops sagax*). Mol. Ecol. **13**(8): 2169–2182. doi:10.1111/j.1365-294X.2004.02229.x. PMID:15245392.

Lenfest Forage Fish Task Force. 2012. Little fish big impact: managing a crucial link in ocean food webs – a report from the Lenfest Forage Fish Task Force. Available from https://www.oceanconservationscience.org/foragefish/files/Little%20Fish,%20Big%20Impact.pdf.

Lewis, L.S., Willmes, M., Barros, A., Crain, P.K., and Hobbs, J.A. 2020. Newly discovered spawning and recruitment of threatened longfin smelt in restored and underexplored tidal wetlands. Ecology, **101**: e02868. doi:10.1002/ecy.2868. PMID:31461779.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics, **25**(16): 2078–2079. doi:10.1093/bioinformatics/btp352. PMID:19505943.

Litz, M.N.C., Emmett, R.L., Bentley, P.J., Claiborne, A.M., and Barceló, C. 2014. Biotic and abiotic factors influencing forage fish and pelagic nekton community in the Columbia River plume (USA) throughout the upwelling season 1999–2009. ICES J. Mar. Sci. **71**(1): 5–18. doi:10.1093/icesjms/fst082.

Mac Nally, R., Thomson, J.R., Kimmerer, W.J., Feyrer, F., Newman, K.B., Sih, A., et al. 2010. Analysis of pelagic species decline in the upper San Francisco Estuary using multivariate autoregressive modeling (MAR). Ecol. Appl. **20**(5): 1417–1430. doi:10.1890/09-1724.1. PMID:20666258.

Maunder, M.N., Deriso, R.B., and Hanson, C.H. 2015. Use of state-space population dynamics models in hypothesis testing: advantages over simple log-linear regressions for modeling survival, illustrated with application to longfin smelt (*Spirinchus thaleichthys*). Fish. Res. **164**: 102–111. doi:10.1016/j.fishres.2014.10.017.

Meisner, J., and Albrechtsen, A. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics, **210**(2): 719–731. doi:10.1534/genetics.118.301336. PMID:30131346.

Miller, M.R., Atwood, T.S., Eames, B.F., Eberhart, J.K., Yan, Y.-L., Postlethwait, J.H., and Johnson, E.A. 2007. RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biol. **8**(6): R105. doi:10.1186/gb-2007-8-6-r105. PMID:17553171.

Moulton, L.L. 1974. Abundance, growth, and spawning of the longfin smelt in Lake Washington. Trans. Am. Fish. Soc. **103**(1): 46–52. doi:10.1577/1548-8659(1974)103<46:AGASOT>2.0.CO;2.

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

Nobriga, M.L., and Rosenfield, J.A. 2016. Population dynamics of an estuarine forage fish: disaggregating forces driving long-term decline of longfin smelt in California's San Francisco Estuary. Trans. Am. Fish. Soc. **145**(1): 44–58. doi:10.1080/00028487.2015.1100136.

Pearcy, W., Fisher, R., Brodeur, R., and Johnson, S. 1985. Effects of the 1983 El Nino on coastal nekton off Oregon and Washington. *In* El Nino North – Nino Effects in the Eastern Subarctic Pacific Ocean. *Edited by* W.S. Wooster and D.L. Fluharty. Washington Sea Grant Program, Seattle, Wash. pp. 188–204.

Pequegnat, J.E., and Butler, J.H. 1982. The biological oceanography of Humboldt Bay. *In* Proceedings of the Humboldt Bay Symposium, Eureka, Calif., 26 March, 1982. *Edited by* C. Toole and C. Diebel. Center for Community Development, Humboldt State University, Arcata, Calif. pp. 39–51.

Penttila, D. 2007. Marine forage fishes in Puget Sound. Puget Sound Nearshore Partnership Report No. 2007-03. Seattle District, U.W Army Corps of Engineers, Seattle, Wash.

Pikitch, E.K., Rountos, K.J., Essington, T.E., Santora, C., Pauly, D., Watson, R., et al. 2014. The global contribution of forage fish to marine fisheries and ecosystems. Fish Fish. **15**(1): 43–64. doi:10.1111/faf.12004.

Purvis, N.H. 1934. History of the Lake Washington canal. Wash. Hist. Quart. **25**(2): 114–127.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. **67**(5): 901.

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Rosenfield, J.A., and Baxter, R.D. 2007. Population dynamics and distribution patterns of longfin smelt in the San Francisco Estuary. Trans. Am. Fish. Soc. **136**(6): 1577–1592. doi:10.1577/T06-148.1.

Sağlam, İ.K., Baumsteiger, J., Smith, M.J., Linares-Casenave, J., Nichols, A.L., O'Rourke, S.M., and Miller, M.R. 2016. Phylogenetics support an ancient common origin of two scientific icons: Devils Hole and Devils Hole pupfish. Mol. Ecol. **25**(16): 3962–3973. doi:10.1111/mec.13732.

Santos, N., Viers, J. Katz, J., and Moyle, P. 2014. Longfin Smelt Range [ds1324]. UC Davis PISCS, California Fish Data and Management Software. California Department of Fish and Wildlife's Biogeographic Information and Observation System. Available from https://wildlife.ca.gov/Data/BIOS.

Sibley, T.H., and Chigbu, P. 1994. Feeding behavior of longfin smelt (*Spirinchus thaleichthys*) may affect water quality and salmon production in Lake Washington. Lake Reserv. Manage. **9**(1): 145–148. doi:10.1080/07438149409354747.

Skinner, J.E. 1962. An historical review of the fish and wildlife resources of the San Francisco Bay Area. Water Projects Branch Report No. 1. The Resources Agency of California, Department of Fish and Game, Water Projects Branch. Available from http://stripersurf.com/california_history.html.

Skotte, L., Korneliussen, T.S., and Albrechtsen, A. 2013. Estimating individual admixture proportions from next generation sequencing data. Genetics, **195**(3): 693–702. doi:10.1534/genetics.113.154138. PMID:24026093.

Sommer, T., Armor, C., Baxter, R., Breuer, R., Brown, L., Chotkowski, M., et al. 2007. The collapse of pelagic fishes in the Upper San Francisco Estuary. Fisheries, **32**(6): 270–277. doi:10.1577/1548-8446(2007)32[270:TCOPFI]2.0.CO;2.

Sutherland, B.J.G., Candy, J., Mohns, K., Cornies, O., Jonsen, K., Le, K., et al. 2021. Population structure of eulachon (*Thaleichthys pacificus*) from Northern California to Alaska using single nucleotide polymorphisms from direct amplicon sequencing. Can. J. Fish. Aquat. Sci. **78**(1): 78–89. doi:10.1139/cjfas-2020-0200.

Thomson, J.R., Kimmerer, W.J., Brown, L.R., Newman, K.B., Mac Nally, R., Bennett, W.A., et al. 2010. Bayesian change point analysis of abundance trends for pelagic fishes in the upper San Francisco Estuary. Ecol. Appl. **20**(5): 1431–1448. doi:10.1890/09-0998.1. PMID:20666259.

USFWS. 1994. Endangered and threatened wildlife and plants: notice of 1-year findings on a petition to list the longfin smelt. US Fish and Wildlife Service. Federal Register: January 6, 1994. 59 FR(4): 869–871.

USFWS. 2012. Notice of 12-month petition finding (longfin smelt). Docket No. FWS-R8-Es-2008-0045. US Fish and Wildlife Service (USFWS), Washington, D.C.

Wilson, G.A., and Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics, **163**(3): 1177–1191. doi:10.1093/genetics/163.3.1177. PMID:12663554.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by 88.255.99.38 on 12/28/22