Genetic Structure of Delta Smelt in the San Francisco Estuary

Kathleen M. Fisch[1,2], Ronald S. Burton[2] & Bernie May[1]

[1]*Genomic Variation Laboratory, Department of Animal Science, UC Davis, Davis, CA, USA*

[2]*Marine Biology Research Division, Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, USA*

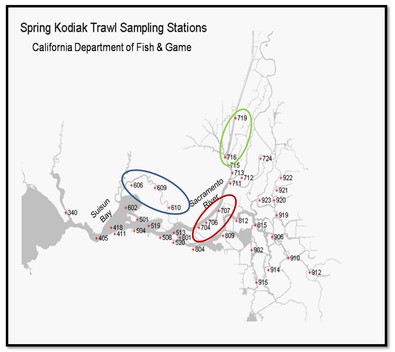

Figure 1. Map of CDFG Spring Kodiak Trawl Sampling Stations. Samples analyzed from stations 606, 609, 610, 704, 706, 707, 716, 719. Other stations were excluded as they did not have enough samples for statistical analysis.

Introduction

Delta smelt, *Hypomesus transpacificus*, are threatened with extinction due to extreme anthropogenic alterations to their ecosystem. They are pelagic planktivorous fish endemic to the upper San Francisco Estuary (SFE), California, USA, a highly urbanized ecosystem affected by the introduction of nonnative species, water diversions, contaminants, and the creation of an extensive levee system, all considered threats to delta smelt (Moyle 2008; Nichols et al. 1986).

Understanding the genetic structure of delta smelt in the SFE is important for directing conservation recommendations to managers. Historically, delta smelt lived in a highly connected delta ecosystem, and consequently may have had greater genetic connectivity before the creation of the extensive levee system. The impacts on the genetic diversity of delta smelt due to the highly altered system in which they live has previously been unknown.

To assess the genetic structure of delta smelt, we conducted a genetic study using sixteen microsatellite DNA markers. We genetically analyzed 353 individuals collected by the California Department of Fish & Game in the 2009 Spring Kodiak Trawl Survey, consisting of delta smelt collected from eight unique geographic sampling locations.

Methods

- Samples (*n*=353) from CDFG 2009 Spring Kodiak Trawl Survey (Figure 1).
  - Genomic DNA extracted with Qiagen DNeasy Tissue Kit.
  - 16 microsatellite markers amplified with PCR (Fisch et al. 2009).
  - PCR products visualized on ABI 3730 DNA Analyzer with LIZ500 internal size standard.
  - Alleles scored with ABI's Genemapper 4.0.
- *Genetic diversity*
  - Number of alleles per locus (A), observed heterozygosity (H_O) & expected heterozygosity (H_E) were calculated using Cervus 3.0 (Kalinowski *et al.* 2007).
  - Allelic richness (A_R) was calculated as a measure of the number of alleles independent of sample size using FSTAT 2.9.3 (Goudet 2001).
  - Tests of Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were conducted using GenePop 3.4 (Raymond & Rousset 1995).
- *Population Differentiation*
  - Population pairwise comparisons of F_{ST} were calculated and tested for statistical significance with 10,000 permutations in Arlequin 3.1 (Excoffier *et al.* 2005).
- *Bayesian cluster analysis*
  - To estimate the number of populations represented by the sampled individuals (*k*), a Bayesian clustering method was implemented in Structure 2.3.3 (Pritchard *et al.* 2000).
  - Assuming admixture and correlated allele frequencies, 25 independent runs assuming k = 1 to 10 were performed with 1×10^6 Markov chain Monte Carlo (MCMC) repetitions and a burn-in period of 5×10^5 using no prior information and then again with prior population information (Hubisz et al. 2009).
  - To determine the most likely value for *k*, the steepest increase of the probability of *k* was measured by plotting the probability of the data [*P*(D)] and the *ad hoc* statistic Δ*k* (Evanno *et al.* 2005).

Results

- **Number of alleles per locus**: A total of 361 alleles were detected among the 16 microsatellite loci. The number of alleles per locus ranged from 6 (*HtrG118*) to 34 (*HtrG119*) (Table 1).
- **Allelic richness**: Ranged from 4.7 to 29 alleles at each locus. Allelic richness was not significantly different among populations (Table 1).
- **Mean H_E**: 0.815 (ranging from 0.238 to 0.956) for the total population (Table 2).
- **HWE**: 4 loci *(HtrG103, HtrG104, HtrG116, HtrG126)* deviated from HWE (P<0.05; Table 2).
- **F_{ST} Values**: low differentiation, significant between 716_719 and other two populations (Table 3).
- **STRUCTURE**: One genetic cluster had highest probability with no prior information (Figure 2) but genetic structure detected when sampling stations used as prior information (Figure 4 & 5).
- **PCA**: Indicates one single cluster when geographic information not included (Figure 3).

Analysis without prior group information

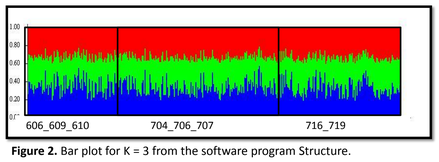

Figure 2. Bar plot for K = 3 from the software program Structure.

Analysis with prior group information

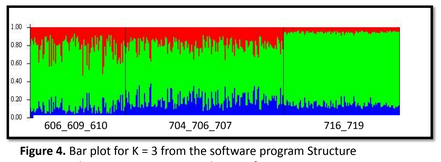

Figure 4. Bar plot for K = 3 from the software program Structure using sampling station as prior population information.

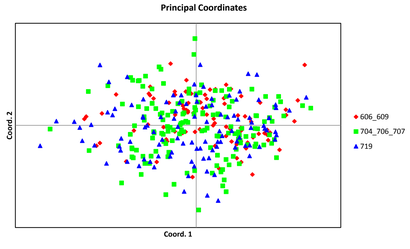

Figure 3. PCA plot indicating one cluster for the three populations.

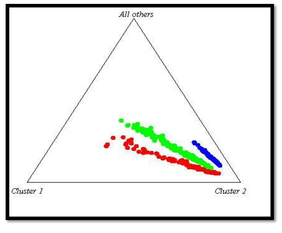

Figure 5. Triangle plot for K = 3 from the software program Structure using sampling station as prior population information.

Table 1 Allelic diversity of delta smelt at 16 microsatellite loci, including locus name, number of alleles per locus (N_a), number of individuals genotyped at each locus (N), number of alleles in each population (A), number of private alleles (Pr) in each population, and allelic richness* (A_R) for each population.

		Overall Population			Population 1				Population 2				Population 3			
		2009 SKT			606_609_610				704_706_707				716_719			
Locus	N_a	N	A	A_R	N	A	Pr	A_R	N	A	Pr	A_R	N	A	Pr	A_R
HtrG103	25	324	25	20	74	22	2	19	143	22		17	107	19	1	19
HtrG104	14	335	14	9.5	89	12	1	8	137	10	1	8.1	109	10	1	8.6
HtrG107	21	278	21	20	54	20	1	16	130	19		14	94	15		16
HtrG109	20	294	20	14	66	15		15	123	17	1	15	104	18	1	15
HtrG114	29	341	29	21	85	23	2	22	150	26	2	22	106	25	1	22
HtrG115	28	336	28	21	88	23	1	22	141	27	4	17	107	20		20
HtrG116	11	310	11	7.8	71	8		9.2	142	11	2	5.3	97	6		8.2
HtrG117	22	302	22	16	67	16	1	17	132	20	3	15	103	18		16
HtrG118	6	335	6	4.7	80	5		5.3	146	6		4.5	109	6		4.8
HtrG119	34	318	34	29	77	31	1	26	143	29	2	28	98	30	1	27
HtrG120	21	336	21	16	88	19	1	14	139	18	1	14	109	17		15
HtrG126	33	312	33	24	71	26	2	26	135	30	3	22	106	25	1	25
HtrG127	33	341	33	26	89	27		28	144	32	1	28	108	31	1	28
HtrG128	21	307	21	17	65	18		16	139	20	1	16	103	19	1	16
HtrG129	14	266	14	10	53	10		10	128	14	4	6.5	85	7		8.8
HtrG131	29	340	29	23	87	25	1	24	149	27	1	23	104	26		24
Average		---	---	**17**	---	---	---	**17**	---	---	---	**16**	---	---	---	**17**

*A_R based on min sample size of 53 diploid individuals

Table 2. Heterozygosity of delta smelt, including locus name, numbers of individuals genotyped (*N*), observed heterozygosity (H_O), expected heterozygosity (H_E), HWE *P* values (*P*), and *F*-statistics within population (F_{IS}).

	Overall Population				Population 1				Population 2				Population 3			
	SKT2009				606_609_610				704_706_707				716_719			
Locus	*N*	H_O	H_E	*P*	F_{IS}	H_O	H_E	*P*	F_{IS}	H_O	H_E	*P*	F_{IS}	H_E	H_O	*P*	F_{IS}
HtrG103	324	0.843	0.895	0.000*	0.058	0.797	0.883	0.013	0.098	0.839	0.898	0.006	0.066	0.898	0.879	0.666	0.022
HtrG104	335	0.576	0.579	0.0005*	0.005	0.607	0.578	0.377	-0.050	0.591	0.593	0.010	0.002	0.566	0.532	0.094	0.060
HtrG107	278	0.723	0.785	0.003	0.079	0.815	0.832	0.514	0.021	0.700	0.790	0.0017*	0.114	0.747	0.702	0.113	0.061
HtrG109	294	0.891	0.886	0.477	-0.006	0.924	0.883	0.147	-0.047	0.903	0.889	0.796	-0.016	0.884	0.856	0.569	0.032
HtrG114	341	0.927	0.940	0.966	0.014	0.906	0.936	0.262	0.033	0.927	0.941	0.366	0.015	0.940	0.943	0.576	-0.004
HtrG115	336	0.940	0.932	0.097	-0.009	0.909	0.923	0.044	0.027	0.965	0.936	0.887	-0.030	0.924	0.935	0.080	-0.012
HtrG116	310	0.639	0.598	0.0024*	-0.068	0.718	0.633	0.214	-0.135	0.690	0.624	0.006	-0.106	0.533	0.505	0.302	0.052
HtrG117	302	0.887	0.917	0.281	0.032	0.881	0.912	0.426	0.034	0.856	0.920	0.111	0.070	0.909	0.932	0.930	-0.026
HtrG118	335	0.266	0.277	0.248	0.042	0.188	0.238	0.019	0.212	0.336	0.300	0.613	-0.118	0.277	0.229	0.040	0.172
HtrG119	318	0.934	0.952	0.140	0.019	0.935	0.956	0.460	0.022	0.944	0.951	0.489	0.007	0.953	0.918	0.569	0.022
HtrG120	336	0.830	0.823	0.382	-0.009	0.818	0.816	0.700	-0.003	0.856	0.827	0.180	-0.036	0.824	0.807	0.369	0.021
HtrG126	312	0.859	0.948	0.000*	0.094	0.859	0.948	0.0008*	0.094	0.874	0.950	0.0003*	0.080	0.945	0.840	0.028	0.112
HtrG127	341	0.933	0.960	0.937	0.029	0.933	0.959	0.400	0.028	0.972	0.960	0.962	-0.012	0.960	0.880	0.092	0.084
HtrG128	307	0.902	0.889	0.703	-0.015	0.939	0.908	0.911	-0.033	0.921	0.886	0.305	-0.040	0.883	0.854	0.531	0.032
HtrG129	266	0.733	0.706	0.273	-0.039	0.736	0.717	0.782	-0.027	0.734	0.710	0.204	-0.035	0.695	0.729	0.752	-0.050
HtrG131	340	0.956	0.948	0.560	-0.008	0.989	0.948	0.575	-0.043	0.940	0.951	0.122	0.012	0.945	0.952	0.094	-0.008
Average		**0.802**	**0.815**		**0.015**	**0.810**	**0.817**	**0.389**	**0.014**	**0.816**	**0.820**	**0.361**	**-0.002**	**0.805**	**0.781**	**0.363**	**0.036**

*Statistically significant at P<0.05

Conclusions

1. **Delta smelt have a high degree of genetic diversity.**
2. **F_{ST} values indicate that samples from Sacramento Deep Water Ship Channel have low, but significant genetic divergence from samples in the lower Sacramento River and Montezuma Slough.**
3. **Using prior information on population groupings, Bayesian analyses using STRUCTURE reveals population differentiation consistent with the F_{ST} analysis. This differentiation is not apparent when this analysis does not include the prior information.**
4. **Further studies using existing samples from earlier collection years should be conducted to elucidate the temporal population structure of delta smelt to determine if the observed divergence is a recent phenomenon due to anthropogenic fragmentation of the Delta or a consequence of delta smelt evolution.**

Table 3. F_{ST} values between sampling stations in the SKT2009 delta smelt population (lower diagonal).

	606_609_610	704_706_707	716_719	FIS
606_609_610	——	NS	*	-0.03647
704_706_707	-0.00465	——	*	0.03634
716_719	0.00646	0.0094	——	-0.02563

* Significant (P<0.05) differentiation is indicated with * (upper diagonal). NS = not significant.

Literature cited

Evanno, G. Regnaut, S., & J. Goudet. 2005. Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology,* 14, 2611-2620.

Excoffier, L., Laval, G. & Schneider. 2005. Arelquin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Fisch, K., Petersen, J., Baerwald, M., Pedroia, J., & B. May. 2009. Characterization of 24 microsatellite loci in delta smelt, Hypomesus transpacificus, and their cross-species amplification in two other smelt species of the Osmeridae family. *Molecular Ecology Resources* 9(1): 405-408.

Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995), available from http://www.unil.ch/izea/softwares/fstat.html.

Hubisz, M., Falush, D., Stephens, M. & J. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322-1332.

Kalinowski, S., Taper, M, & Marshall, T. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**: 1099-1006.

Moyle, P.B. 2008. The future of fish in response to large-scale change in the San Francisco Estuary, California. *American Fisheries Society Symposium* 64: 1-18.

Nichols, F. , J. Cloern, S. Luoma, & D. Peterson. 1986. The modification of an estuary. *Science* 231:567–573.

Pritchard J., Stephens M., & Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Raymond, M. & F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumeniscism. *Journal of Heredity*, 86, 246-249.

Acknowledgments

We thank Brian Mahardja and Jordana Henderson for laboratory support and Randy Baxter, Julio Adib-Samii, Erin Gleason and Kelly Souza for providing delta smelt samples. The USFWS supported this research through a CESU agreement with UC Davis (Agreement no. 813327J011).

For further information

Please contact *kmfisch@ucdavis.edu*. More information on this and related projects can be obtained at http://www.genome-lab.ucdavis.edu.