Spatial and ecological structure of the sockeye salmon population *Oncorhynchus nerka* of Kuril Lake

Anastasia Khrustaleva*

Institute of Gene Biology Russian Academy of Sciences (FSBIS IGB RAS), group of bioinformatics, 34/5 Vavilova Street, Moscow, 119334, Russia

Abstract. This study examines the genetic structure of sockeye salmon populations in Kuril lake basin, focusing on the allelic frequencies of SNP loci that potentially carry an adaptive load. Our results reveal the division of sockeye salmon populations into three distinct groups: early river sockeye salmon and two groups of littoral sockeye salmon from the southern and northeastern shores of the lake. This division is supported by the genetic data, which shows that the populations are genetically distinct and have distinct allelic frequencies. Additionally, our analysis reveals that the neutral structure of the population is rather shallow, with only minor differences found between the two ecotypes of sockeye salmon - stream and littoral. These findings suggest that the genetic differences between the populations are primarily driven by adaptive processes, rather than neutral genetic drift.

1 Introduction

Sockeye salmon is a commercially important species and is one of the three most numerous species of Pacific salmon in the Russian Far East. It is characterized by an anadromous lifestyle (enters rivers from the sea to spawn) with a long freshwater period, monocyclicity (all fish die after spawning), high homing (instinct to return to their native water body) and is represented throughout its range by many geographical populations (or metapopulations, i.e. systems of relatively isolated subpopulations inhabiting a river basin). Kuril Lake in the Ozernaya river basin is a spawning and feeding reservoir for the largest sockeye salmon stock on the Asian Pacific coast, the number of which in some years exceeded 20 million. [1]. The spawning run of Ozernovskaya sockeye salmon usually begins in late May and ends in late October – early November [1]. Sockeye salmon gradually fills the spawning areas in the lake basin, moving from north to south clockwise, with three quarters of the spawning areas in littoral areas, and only a quarter of the area in river spawning areas. In the first half of August, sockeye salmon spawners enter the tributaries of the lake; a little later, in the second half of August, the littoral spawning grounds are filled. Depending on the preference for one or another spawning biotope, sockeye salmon are usually distinguished into stream and littoral ecotypes, which differ phenotypically: small, faster

---

* Corresponding author: mailfed@mail.ru

fish are more successful in shallow spawning grounds in streams and small rivers, large and tall-bodied fish are more successful in deeper-water littoral areas [2]. Such a complex spatial and ecological structure of the sockeye salmon herd from Kuril Lake in combination with strict homing can contribute to the genetic divergence of subpopulations of different biotopes and lake shores.

The Kuril Lake basin has in recent years represented a steady model for organizing effective environmental management. A KamchatNIRO observation post operates on Kuril Lake, conducting regular fish surveys and annual population assessments to monitor the Ozernovskaya sockeye salmon stock. This comprehensive monitoring program ensures a high level of management for this important salmon population. Additionally, a fishing regulation system has been established based on optimal indicators to ensure sufficient passage of spawning salmon to the lake's spawning grounds. However, in some years there were significant changes in the dynamics of the spawning run of sockeye salmon spawners in large tributaries of the lake, most likely due to uneven fishing pressure on the components of this population system, as well as a result of the action of natural factors - density-dependent, trophic, global or local climatic and hydrological [3,4]. And yet, the main reason, apparently, is the non-selectivity of fishing due to insufficient or ignored information about the population structure and ecological-temporal differentiation of the Kuril Lake population in planning and organizing fishing. Excessive fishing pressure on individual components of a population system can lead to a reduction in their effective numbers or to their complete elimination, and, accordingly, to a decrease in ecological and genetic diversity, adaptive potential and sustainability of the system as a whole.

In this regard, the purpose of this work was to study the neutral and adaptive genetic structure of the sockeye salmon population in Kuril Lake and to identify spatial, ecological and temporal subdivision of the metapopulation of the Kuril lake using analysis of the variability of two types of molecular DNA markers - single nucleotide polymorphism (SNP) and microsatellite loci.

2 Materials and methods

This paper presents the results of a meta-analysis of previously published data on polymorphism of 45 SNP loci [5,6] and 13 microsatellite loci [7] of sockeye salmon from the River Ozernaya. The final dataset included 19 samples of sockeye salmon collected at 12 sockeye salmon spawning grounds in the Kuril Lake Basin (Fig. 1, Table 1,2).

Statistical analysis was carried out in a general manner for both types of markers: expected heterozygosity and allelic diversity in the samples were assessed, and statistical approaches of multivariate analysis were universally used for both types of markers. Probabilities of population bottleneck events were estimated for 27 neutral SNP loci in Bottleneck 1.2.02 using three tests: the sign test, the standardized difference test, and the Wilcoxon infinite allele model (IAM) test. A bottleneck was considered probable if all three tests were significant. To check bottleneck passage using microsatellite data, the M-ratio was calculated for each locus using the formula $M = k/r$, where k is the number of alleles per locus, r is the range of variation in allele sizes, expressed in the number of repeats, as described previously [8], while the average M-coefficients for model and natural populations that are known to have experienced significant declines in their effective numbers in the past should not exceed 0.68. The calculation of the genetic distances of Cavalli-Sforza and Nei was carried out in the R-package Rphilip; the phylogenetic network was built using the Neighbor-Net algorithm in the R-library phangorn, as described previously [6]. Principal component analysis (Principal Component Analysis, PCA) was carried out using the R libraries factoextra [9] and FactoMineR [10], and coordinate analysis (Principal Coordinate Analysis, PCoA) was carried out using hierfstat [11].

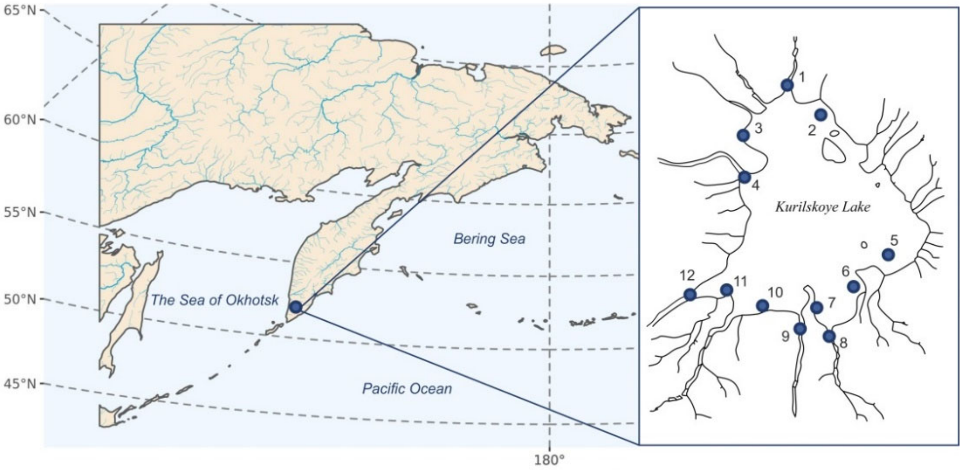

Fig. 1. Scheme map of the material collection area. 1 − Vychenkia R., 2 − Bay Northern Far, 3 − Bay Northern Near, 4 − Bay Ozernaya, 5 − Bay Oladochnaya, 6 − Bay Gavryushka, 7 − Bay Gavryushka, Cape Tugumynk, 8 − Gavryushka R., 9 – Kirushutk R., 10 − Bay Hakitzin, 11 − Bay South, 12 – Etamynk R.

Table 1. Characteristics of sockeye salmon samples from Kuril Lake and their genetic indicators, calculated from the results of the analysis of polymorphism of 13 microsatellite loci according to Varnavskaya, 2006.

Spawning ground	**Coast**	**Designation of selection**	***n***	**Date of catch**	**Description of the sample**	***He(SD)***	***n_a(SD)***	***M-ratio***
R. Vychenkia	north	KOvych	100	28.07.2000	River early running	0.652 (0.222)	11.46 (9.38)	0.686
Bay Northern Far	north	KOsevF	54	30.08.2000	Intertidal mid-season spawning	0.675 (0.208)	9.46 (5.98)	0.683
Bay Northern Near	north	KOsevN	104	30.08.2000	Intertidal mid-season spawning	0.679 (0.192)	10.77 (6.6)	0.636*
Bay Ozernaya	north-west	KOoz	48	09.10.2000	Littoral, late spawning	0.662 (0.214)	9.62 (7.3)	0.588*
Bay Oladochnaya	south-east	KOol	92	08.10.2000	Littoral, late spawning	0.676 (0.186)	10.69 (6.3)	0.629*
Bay Gavryushka	south	KOgav	88	17.08.2000	Intertidal mid-season spawning	0.668 (0.214)	11.38 (6.85)	0.685
Bay Gavryushka, Cape Tugumynk	south	KOgav2	54	18.08.2000	Intertidal mid-season spawning	0.669 (0.203)	9.38 (5.41)	0.645*
R. Gavryushka	south	KOgav R	55	02.08.2000	River early running	0.664 (0.187)	9.38 (5.66)	0.655*

R. Kirushutk	south	KOkir	120	31.07.2000	River early running	0.651 (0.224)	11.92 (7.85)	0.672*
Bay Hakitzin	south	KOhak	53	31.08.2000	Intertidal mid-season spawning	0.64 (0.197)	8.15 (4.88)	0.644*
Bay South	south	KOyuz	100	08.10.2000	Littoral, late spawning	0.663 (0.217)	11.38 (8.82)	0.696
R. Etamynk	south	KOet	55	05.08.2000	River early running	0.664 (0.213)	9.92 (5.2)	0.719

Note: *M-ratio* - M-ratio according to Garza, Williamson, 2001; An asterisk (*) indicates *M*-ratio estimates that indicate a probable "bottleneck" in the population. Here and in Table 2: *n* – number of fish in the sample, specimens; *He* – average expected heterozygosity, n_a – average number of alleles per locus, *SD* – standard deviation

Table 2. Characteristics of sockeye salmon samples from Kuril Lake, estimates of intrapopulation diversity and results of bottleneck tests, calculated from allelic frequencies of 45 single nucleotide polymorphism loci according to Habicht et al., 2010.

Place of catch	**Designation of selection**	***n***	**Date of catch**	**Description of the sample**	***He(SD)***	n_a***(SD)***	p_{sign}	p_{stdv}	p_W
R. Vychenkia	KOvych	96	28.07.2000	River early running	0.236 (0.187)	1.9 (0.3)	0.022*	0.018*	0.063
Bay Northern Far	KOsev	50	26.08.2002	Intertidal mid-season spawning	0.248 (0.19)	1.88 (0.33)	0.006**	0.009**	0.032*
Bay Oladochnaya	KOol	50	08.10.2000	Littoral, late spawning	0.243 (0.191)	1.8 (0.41)	0.012*	0.006**	0.013*
Bay Gavryushka	KOgav	50	25.08.2002	Intertidal mid-season spawning	0.249 (0.189)	1.85 (0.36)	0.02*	0.01*	0.043*
R. Kirushutk	KOkir	49	31.07.2000	River early running	0.249 (0.198)	1.88 (0.33)	0.12	0.025*	0.06
Bay Khakytsin	KOhak	50	31.08.2002	Intertidal mid-season spawning	0.236 (0.188)	1.85 (0.36)	0.176	0.031*	0.073
R. Etamynk	KOet	127	12.08.2002 21.08.1990 28.09.1990	Mixed sampling, river and littoral mid and late spawning periods	0.248 (0.194)	1.93 (0.27)	0.011*	0.011*	0.035*

Note: Significance levels in bottleneck tests using the following criteria: *psign* – signs, *pstdv* – standardized differences, *pW* – Wilcoxon's, * – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$.

3 Results and discussion

Based on the results of the analysis of intrapopulation genetic diversity, any patterns in the distribution of estimates of allelic diversity and expected heterozygosity, calculated for two types of markers, in sockeye salmon populations from Kuril Lake fish reproducing on different spawning grounds were not observed. We can only say unequivocally that the expected heterozygosity was slightly lower according to the estimates obtained for both markers in the sample of sockeye salmon from the River Vychenkia (northern coast). It is also premature to draw any conclusions about the recent decline in the numbers of individual populations based on the results of bottleneck tests. Conflicting estimates were obtained for different types of markers, and only for the population of late-spawning littoral sockeye salmon in Bay Oladochnaya (latest spawning in October) suggests a recent passage of the "neck of the bottle".

Using PCA analysis of allelic frequencies of 45 SNP loci, samples were ordi- nated in the space of principal components (Fig. 2a). The relative position of sockeye salmon samples from different spawning grounds in the Kuril Lake basin on the plane of the first two principal components (explaining in total about 50% of the variance in allelic frequencies of SNP loci) indicates the existence within the lake of three more or less separate groups of sockeye salmon populations. This is a group of populations of early river sockeye salmon and 2 groups of populations of littoral sockeye salmon on the southern and northeastern coasts of the lake. Moreover, if the first component (PC1) differentiates the selected groups well, then the second component (PC2) shows a connection with the latitude of the spawning area. Thus, if we consider separately the clusters of early river and late littoral sockeye salmon, then the populations of the northern coast of the lake are in the upper part of the graph, and the southern coast is in the lower part.

The topology of the split network also indicates the proximity of two river samples of early-spawning sockeye salmon from the Kirushutk and Vychenkiya Rivers (forming a stable common clade) (Fig. 2b). The middle part of the network is occupied by samples of southern littoral spawning grounds. The largest split is observed between them and samples of sockeye salmon spawning in the littoral zone in the bays of the northeastern coast of the lake (Northern Dalnyaya and Oladochnaya Bays). Late spawning sockeye salmon from bay. Oladochnaya was the most different from the rest. These differences only intensified when the network was rebuilt based on distances calculated from the allelic frequencies of presumably neutral loci (27 SNPs) (Fig. 2c), which likely confirms the hypothesis of a decline in the size of this population in the relatively recent past. At the same time, when the network is rebuilt, the division into clusters is completely lost, which may indicate the primary role of adaptive processes in the formation of ecotypes and division into spawning areas of sockeye salmon in the Kuril Lake basin. In confirmation of the above, among the SNPs that made the greatest contribution to the first principal component (PC1), responsible for separating the samples of sockeye salmon from Kuril Lake into three groups (including littoral and riverine ecotypes) were *MHC2_190v2, MHC2_251v2, One_zP3b* and *GPH-414*, also considered as candidate loci for selection in a number of North American and Asian sockeye salmon populations [8,12,13].

The structure of clusters in PCA diagrams and the topology of the phylogenetic network correlate with data on the time of filling of spawning grounds in the Kuril Lake basin and geographic variability of spawning timing in the lake basin. Thus, according to literary data, the spawning of sockeye salmon in the lake basin. Kuril is divided into early and late, and its timing usually differs on the spawning grounds of the northern and southern coasts of the lake. The first breeders appear on nests in the northern part of the lake and in the River Vychenkia (spawning in these areas occurs from the end of July to the end of November), somewhat later, in mid-August, spawning begins in the southern part of the lake and in the River Etamynk (and continues here until the end of January) [1]. Studies using tagging and analysis of polymorphism of microsatellite loci have shown that spawning in the rivers of the northern part of the lake (Vychenkiya River) is earlier (July - mid-August), in the rivers and littoral zone of the eastern part of the lake (Oladochnaya Bay) it occurs in August–September, and on the spawning grounds of the southern and western coasts of the lake (the Rivers Gavryushka, Kirushutk, Khakytsyn, Etamynk) - in mid-August - late September, October. The latest spawning (from late September to February) is observed in small springs in the river basin. Middle Khakytsin, flowing into the Kuril Lake from the south, and in the Etamynk Lake [14].

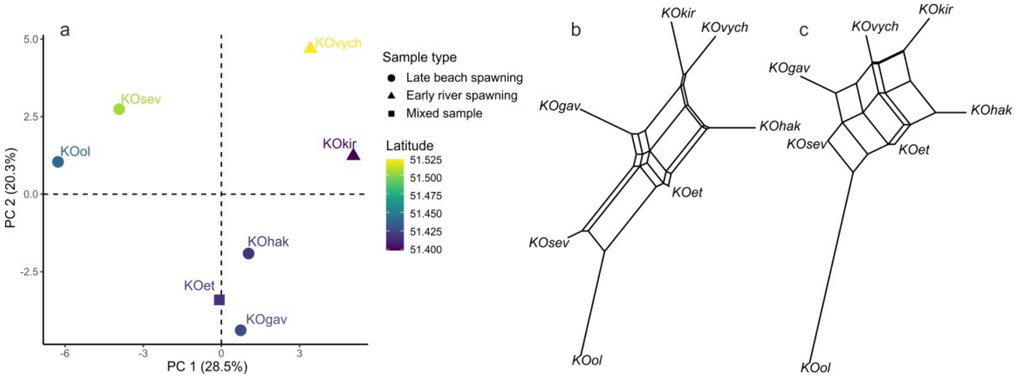

Fig. 2. Ordination of sockeye salmon samples from the Kuril Lake in the space of the first two principal components (PC) based on the results of analysis of allelic frequencies of 45 SNP loci (a): values in parentheses indicate the contribution of each component (% of total variance). Phylogenetic Neighbor-Net networks constructed from chordal distances calculated from allelic frequencies of 45 SNP loci (b) and 27 putatively neutral SNP loci (c). For sample designations see Table 2.

Analysis of the frequencies of 13 microsatellite loci revealed no differences either between northern and southern spawners, or between intertidal populations differing in spawning time (mid-term spawning and October spawning) (Fig. 3a, b). On the plane of the first two main coordinates, only a more or less pronounced differentiation into two clusters is noticeable - river and littoral populations. At the same time, the 95% confidence intervals of both groups overlapped significantly, and the sample of littoral sockeye salmon Gavryushka Bay fell into the cluster of river sockeye salmon (it is quite possible that the sample from the mouth zone could contain an admixture of sockeye salmon from the Gavryushka River settling before entering the tributary of the lake). The same clusters are also identified on the split phylogenetic network. There are no other patterns in the location of samples on the PCoA diagram and the phylogenetic network.

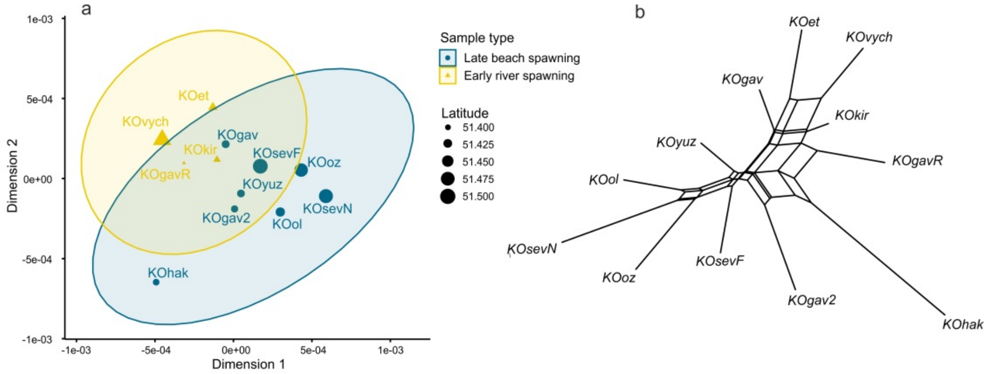

Fig. 3. Location of sockeye salmon samples from Kuril Lake in the space of the first two principal coordinates (Dimensions) based on the results of polymorphism analysis of 13 microsatellite loci (a), ellipses – 95% confidence intervals for groups. Phylogenetic Neighbor-Net networks built using Nei distances (b). For sample designations, see Table 1.

Previously, based on the results of the analysis of polymorphism of 14 microsatellite loci of sockeye salmon (13 used in this work and Ots3), it was demonstrated that the differences between the riverine and littoral ecotypes of sockeye salmon from Kuril Lake (hierarchical analysis of molecular variation (AMOVA)) exceeded the differences between samples of one ecotype [7], however, significant differences between ecotypes were traced only at 5 loci out of 14. In addition, heterogeneity of allelic frequencies was revealed between samples of early river sockeye salmon and middle littoral sockeye salmon (according to 5 loci) and late stages of migration (by 12 loci), as well as between two temporal forms of littoral sockeye salmon (by 8 loci out of 14) [14]. However, no final tests were carried out for all loci in total, so these results are premature to consider justified. According to more modern studies by the same authors, low differentiation was noted between samples of sockeye salmon of early and late forms of spawning in Kuril Lake (Fst = 0.009) based on the frequencies of seven microsatellite loci [15].

The low resolution of microsatellite markers at the intrapopulation level is likely due to the fact that divergence at microsatellite loci is mainly due to genetic drift and mutations, so differences between samples reflect only the duration and degree of isolation between populations. It is obvious that within the same reservoir of reproduction, the isolation between individual spawning grounds is much lower than between neighboring river basins. Nevertheless, the reproductive success of migrants of another ecotype is significantly limited due to sexual selection, adaptive advantage of native populations, and selective grazing by predators. All these processes lead to isolation through adaptation. The differences between sockeye salmon ecotypes in the same lake basin identified in the present study using microsatellite data indicate that neutral markers are capable of detecting not only allopatric or sympatric isolation between sockeye salmon populations, but also isolation through adaptation at the level of spawning habitat types. In contrast to neutral genetic markers, SNPs, including those localized in coding or regulatory regions of the genome, can characterize both neutral demographic processes in populations and adaptive processes, and should be markers of choice when solving problems of identifying the spatial and ecological structure of individual populations. Previously, using the example of American sockeye salmon, their advantages for differentiating intrapopulation groups, temporal races and ecotypes in river populations of the Bristol Bay and Gulf of Alaska basins were convincingly demonstrated [12,13].

4 Conclusion

Meta-analysis of data on polymorphism of microsatellite loci and SNP loci of sockeye salmon reproducing on different spawning grounds in the Kuril Lake basin., suggests the existence within a given stock of subdivision into more or less separate temporal, ecological and/or geographical groupings. According to our results and literature data for both types of markers, differences were found between two ecotypes of sockeye salmon - stream and littoral, the spawning of which is confined to river and lake biotopes. If SNP markers reveal the divergence of subpopulations on the northern and southern shores of the lake, as well as groups that begin to reproduce at different times and of different ecotypes, then the neutral structure of this population is rather weak. Genetic differentiation of sockeye salmon populations within the Kuril Lake basin, identified with the help of potentially selective markers, indicates adaptive processes prevailing in the formation of the structure of this stock, which is recommended to be taken into account when predicting the number and rational exploitation of Osernaya River sockeye salmon, determining the dates of passing days and assessing fishing pressure on individual groups.

The study was supported by a grant from the Russian Science Foundation No. 23-24-00307 (https://rscf.ru/project/23-24-00307/).

References

1. A. V. Bugaev and V. A. Dubynin, in *Chron. Kronotsky Reserv.* (Kronotsky State Nature Reserve, Elizovo, 2019), pp. 64–73
2. W. A. Larson, T. H. Dann, M. T. Limborg, G. J. McKinney, J. E. Seeb, and L. W. Seeb, Mol. Ecol. **28,** 2254 (2019)
3. S. A. Kolchin, E. V. Volkova, L. V. Pokrovskaya, and A. V. Zavadskaya, Nat. Conserv. Res. **6,** 53 (2021)
4. V. A. Dubynin, E. V. Lepskaya, and E. A. Kirillova, Res. Aquat. Biol. Resour. Kamchatka North-West Part Pacific Ocean **63,** 73 (2021)
5. C. Habicht, L. W. Seeb, K. W. Myers, E. V. Farley, and J. E. Seeb, Trans. Am. Fish. Soc. **139,** 1171 (2010)
6. A. M. Khrustaleva, J. Ichthyol. **63,** 90 (2023)
7. N. V. Varnavskaya, *Geneticheskaya Differentsiatsiya Populyatsii Tikhookeanskikh Lososei* (Publishing house KamchatNIRO, Petropavlovsk-Kamchatsky, 2006)
8. A. M. Khrustaleva, Divers. **15,** 853 (2023)
9. A. Kassambara and F. Mundt, Agricultural Sciences, **12(3)** 159-180 (2020)
10. F. Husson, J. Josse, S. Le, J. Mazet, *FactoMineR'. Package FactorMineR* 100 (2020)
11. J. Goudet and T. Jombart, J Goudet, T Jombart - R package version (2020)
12. M. W. Ackerman, C. Habicht, and L. W. Seeb, Trans. Am. Fish. Soc. **140,** 865 (2011)
13. D. Gomez-Uchida, J. E. Seeb, M. J. Smith, C. Habicht, T. P. Quinn, and L. W. Seeb, BMC Evol. Biol. **11,** 48 (2011)
14. O. A. Pilganchuk and N. V. Varnavskaya, Research of Aquatic Biological Resources of Kamchatka and the North-Western Pacific Ocean **17,** 92 (2010)
15. O. A. Pilganchuk, U. O. Muravskaya, V. V. Savenkov, A. D. Denisenko, and D. A. Zelenina, in Salmon Fish Biology, Reproduction, Fishery. *Materials of the All-Russian Scientific and Practical Conference* (All-Russian Research Institute of Fisheries and Oceanography, Murmansk, 2023), pp. 324–329.