



OPEN ACCESS

EDITED BY

Rachid Mentag,
National Institute for Agricultural
Research, Morocco

REVIEWED BY

Gulab Khedkar,
Dr. Babasaheb Ambedkar Marathwada
University, India

*CORRESPONDENCE

Nicholas Tucker
✉ ntucker@fui.edu

RECEIVED 28 June 2023

ACCEPTED 25 August 2023

PUBLISHED 14 September 2023

CITATION

Timm LE, Tucker N, Rix A,
LaBua S, López JA, Boswell KM
and Glass JR (2023) The untapped
potential of seascape genomics in
the North Pacific.
Front. Conserv. Sci. 4:1249551.
doi: 10.3389/fcosc.2023.1249551

COPYRIGHT

© 2023 Timm, Tucker, Rix, LaBua, López,
Boswell and Glass. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

The untapped potential of seascape genomics in the North Pacific

Laura E. Timm^{1,2}, Nicholas Tucker^{3*}, Anna Rix^{4,5},
Savannah LaBua^{2,3}, J. Andrés López², Kevin M Boswell³
and Jessica R Glass²

¹Genetics Program, Auke Bay Laboratories, Alaska Fisheries Science Center, Juneau, AK, United States, ²College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, Fairbanks, AK, United States, ³Department of Biological Sciences, Florida International University, Miami, FL, United States, ⁴Department of Geography, University of Durham, Durham, United Kingdom, ⁵College of Natural Science and Mathematics, University of Alaska Fairbanks, Fairbanks, AK, United States

Seascape genomics provides a powerful framework to evaluate the presence and strength of environmental pressures on marine organisms, as well as to forecast long term species stability under various perturbations. In the highly productive North Pacific, forage fishes, key trophic links across ecosystems, are also contending with a rapidly warming climate and a litany of associated oceanographic changes (e.g., changes in salinity, dissolved oxygen, pH, primary production, etc.). These changes can place substantial selective pressures on populations over space and time. While several population genomics studies have targeted forage fishes in the North Pacific, none have formally analyzed the interactions between genotype and environment. However, when population genomics studies provide collection location information and other critical data, it is possible to supplement a published genomic dataset with environmental data from existing public databases and perform “*post hoc* seascape genomics” analyses. In reviewing the literature, we find pertinent metadata (dates and locations of sample collection) are rarely provided. We identify specific factors that may impede the application of seascape genomics methods in the North Pacific. Finally, we present an approach for supplementing data in a reproducible way to allow for *post hoc* seascape genomics analysis, in instances when metadata are reported. Overall, our goal is to demonstrate – via literature review – the utility and importance of seascape genomics to understanding the long term health of forage fish species in the North Pacific.

KEYWORDS

forage fishes, Arctic, genotype-environment association analyses, population genomics, metadata, fisheries management

1 Introduction

Seascape genomics is a means of investigating the interactions between the genome of an organism and the marine environment. While initially developed to elucidate the role of ocean currents in shaping spatial genetic structure (Galindo et al., 2006; Liggins et al., 2013), technological advancements in sequencing have empowered the field of seascape genomics to focus on identifying and characterizing genomic regions that are adaptive to different environmental conditions in marine species (see Galindo et al., 2006; Liggins et al., 2013; Selkoe et al., 2016 for a thorough review of the theory and practices). The insights that can be gained from genotype-environment association (GEA) approaches are of high value to perturbed or changing environments, where the adaptive potential of affected species is critical to maintaining ecosystem function.

Of the world's oceans, the Pacific Ocean is unique in that it is the largest and contains some of the oldest deep waters. The North Pacific, in particular, is a marine region of high primary productivity (Hinga, 1985), wherein forage fishes, the species linking trophic levels, serve a critical role in the ecosystem and are closely related to ecosystem health. Broadly, the category of “forage fishes” is informally defined to include small pelagic species that may experience abundance cycles of boom-and-bust (National Marine Ecosystem Status, 2020). In its forage fishes fisheries management plan (FFMP), the National Oceanic and Atmospheric Administration's Alaska Fisheries Science Center (AFSC), defines twelve taxonomic groups within Teleostei, ranging from species such as Pacific herring (*C. lupea pallasii*) and Arctic cod (*Boreogadus saida*) to entire clades like gunnells (*Pholidae*; Ormseth and Yasumiishi, 2019).

Forage fishes are particularly vulnerable to environmental changes associated with climate change (Hollowed et al., 2012; Gobler et al., 2018), including deleterious effects on gene diversity. This impact is particularly pronounced in populations that experience heavy fishing pressures (Petrou et al., 2021). As the North Pacific remains susceptible to some of the most drastic perturbations associated with climate change, sustained ecosystem productivity in this region is dependent on the long-term stability of forage fishes. This is especially true given the shifts of most benthic fish species to pelagic depths in response to climate change (Petrik et al., 2020). Suitability of new pelagic depths is directly linked to restructuring of forage fish populations (Hollowed et al., 2012).

As fisheries managers in the North Pacific contend with the impacts of climate change on their focal taxa, seascape genomics has the potential to address the extent to which species may or may not cope with changing ocean conditions based on their ecological and evolutionary history. Seascape genomics' focus on understanding environmental selection pressures exerted on species and populations leads to the identification of genomic sites or larger chromosomal regions adapting to these pressures and a characterization of overall adaptive potential in the species. Both commercially important species and species that require conservation efforts rely on forage fishes as critical prey items, so the application of seascape genomics to understand

the vulnerability of forage fishes to climate change is especially prescient.

Here, we present the requirements for impactful genome-environment analyses in the marine environment, specifically forage fishes in the North Pacific Ocean, critical species in a region of high biological productivity and lucrative commercial fisheries. Given the deficiency of available data on the adaptive potential of forage fishes, we aim to 1) assess the overall state of the seascape and population genomics research focused on forage fishes in the North Pacific; 2) describe the publicly available resources that exist to empower seascape genomics for fisheries management in the region; and 3) identify opportunities to expand seascape genomics analyses in the region using guides we produced to enable *post hoc* seascape genomic analyses when genetic samples were collected without environmental data.

2 A paucity of genomic data

To quantify the extent to which seascape genomics tools have been utilized on forage fishes in the North Pacific, we conducted a literature search on Google Scholar. Articles were identified via searches that included geographic region (search terms: “North Pacific” and “Arctic”) and method (search terms: “seascape genomics” and “population genomics”) before filtering by species (search terms: “forage fishes” and specific common names found in the FFMP). Finally, articles were removed from the literature set for insufficient sample sizes ($n < 5$ individuals across more than 50% of collection regions) or lack of geographic coordinates.

One of the most glaring gaps observed in our initial literature search was a lack of geographic coordinates or, in some cases, localities associated with samples: nine studies were initially identified, but only five reported collection coordinates (55.56%). While we originally intended to run *post hoc* seascape genomic analyses to synthesize the results from our literature review (see section “Opportunities for *Post Hoc* Extension”) the small number of papers that met our criteria and the spatiotemporal disconnect between them prevented us from conducting such analyses. We identified four studies that met our criteria for inclusion, all of which were published within the last three years (Table 1). These studies targeted Pacific herring *Clupea pallasii* (Orlova et al., 2021; Petrou et al., 2021), polar/Arctic cod *Boreogadus saida* (Quintela et al., 2021), and Japanese anchovy *Engraulis japonicus* (Zhang et al., 2020), representing a fraction of the taxonomic diversity of forage fishes in the North Pacific as defined by AFSC (>50 species; 50 CFR Part 679). Of the four studies we reviewed, one lacked sufficient collection date information, providing only the year of sample collection, and sample collection depth was absent from all four studies. Unsurprisingly, given the dearth of seascape genomics studies in the North Pacific, entire geographic regions were unrepresented, including the Gulf of Alaska, the Aleutian Islands, and the Bering Sea (Figure 1). Our findings highlight the scarcity of suitable datasets for conducting seascape genomics in one of the most productive, threatened, and heavily researched regions of the world.

TABLE 1 Summary of population genomics studies targeting forage fishes in the North Pacific, including genomic library preparation method, sample sizes, and geographic regions.

Article	Target Species	Library Prep Method	Geographic Regions	Sample Sizes	No. of Adaptive SNPs Identified	No. of Neutral SNPs Identified	Collection Coordinates Completeness	Collection Date Completeness
Zhang et al., 2020	Japanese anchovy <i>Engraulis japonicus</i>	RADseq	Bohai Sea	100	6	28,776	100%	100% (day, month, year)
			North Yellow Sea	80				
			North East China Sea	60				
			Taiwan Coastal Waters	32				
			Japan Sea	37				
			Pacific side of Japan	80				
Quintela et al., 2021	polar/Arctic cod <i>Boreogadus saida</i>	RADseq to develop SNP panel	East Siberian Sea	41	116 (targeted)	0	100%	0% (year only)
			Laptev Sea	59				
			Kara Sea	80				
			Barents Sea	78				
			Svalbard Area	999				
Orlova et al., 2021	Pacific herring <i>Clupea pallasii</i>	ddRADseq	Bolshoy Vilyuv Lake	5	Not tested	192,433 (may include adaptive SNPs)	100%	100% (month, year)
			Ainskoe Lake	5				
			Nerpiche Lake	5				
			Sea of Okhotsk	8				
			Sea of Japan	10				
			Bering Sea	13				
			Kara Sea	3				
			Kuril Islands	5				
Petrou et al., 2021	Pacific herring <i>Clupea pallasii</i>	RADseq	Washington	315	620	6,098	100% (available in Petrou et al. (2021) supplementary materials)	100% (day, month, year; available in Petrou et al. (2021) supplementary materials)
			British Columbia	664				
			Alaska	125				

Collection data completeness indicates the percent of samples with collection coordinates (as opposed to general localities) and collection dates (month and year are required for supplementing genomic data with environmental data, so any dates lacking either are deemed to be incomplete).

3 Repositories of environmental data

One of the most notable challenges in correlating genetic samples to environmental data is the lack of *in situ* environmental data collection. However, databases such as WorldClim and the World Ocean Database provide access to historical and current environmental data from across the world's oceans. Assembled and maintained by the European Environment Agency, WorldClim (worldclim.org) provides high resolution global climate data at the scale of 1km, from 1970 to 2000 (Fick and Hijmans, 2017). The World Ocean Database (WOD) is hosted by NOAA's National Centers for Environmental Information (Boyer et al., 2016). It enables targeted searches through an online portal, with search criteria including geographic coordinates, dates, environmental variables of interest, etc.

The myriad research interests in the North Pacific, driven by NOAA's AFSC, have resulted in several publicly available, regional environmental datasets. The Alaska Ocean Observing System (AOOS), established in 2003, is part of the US regional network representing NOAA's Integrated Ocean Observing System. The AOOS focuses on providing ocean data to fill knowledge gaps and improve our understanding of the ecosystem while hosting a public data portal that allows users to freely access environmental data from Alaskan coastlines (portal.aos.org).

Submarine physical structures also mediate the distribution of certain forage fishes. Since the 1980s, ShoreZone (shorezone.org), another NOAA initiative, has compiled images to generate maps and datasets of biological and geological attributes along the coasts of Alaska, British Columbia, Washington, and Oregon. Decades of data collection associated with "The Seward Line", a transect

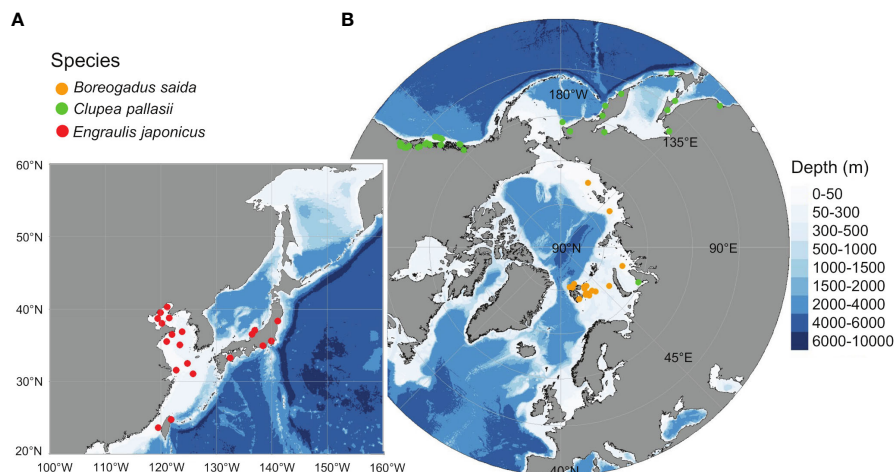


FIGURE 1

Geographic span of the four research articles reviewed here: (A) Zhang et al. (2020), Japanese anchovy *Engraulis japonicus*; and (B) Orlova et al. (2021) and Petrou et al. (2021), targeting Pacific herring *Clupea pallasii*; Quintela et al. (2021), targeting polar/Arctic cod *Boreogadus saida*.

extending from Seward, Alaska to the edge of the continental shelf, have been compiled by the Northern Gulf of Alaska Long Term Ecological Research group (NGA LTER; nga.lternet.edu). Environmental data associated with the Seward Line includes physical hydrography, temperature, salinity, chlorophyll, and nutrients (e.g., nitrate, nitrite, phosphorus, and silicic acid).

4 Opportunities for post hoc extension

The investment that has gone into generating genomic data (Zhang et al., 2020; Orlova et al., 2021; Petrou et al., 2021; Quintela et al., 2021) and recording environmental observations in the North Pacific (AOOS, WOD, WorldClim) offers a novel opportunity to investigate GEAs *post hoc*. When investigating GEAs is not the explicit objective of a population genomics study, the collection of real-time environmental measures is frequently infeasible. However, seascape genomics studies require robust environmental data. While some forms of data necessitate high costs and vessel time (e.g. CTD and YSI profiles), researchers can bypass these barriers and undertake *post hoc* seascape genomics studies by supplementing existing population genomics datasets with publicly available environmental data.

To empower *post hoc* GEA research using existing resources, we designed a series of guides for leveraging publicly available environmental databases, including AOOS (see https://github.com/SimplySav101/Seascape-Genomics-of-North-Pacific-Forage-Fishes/blob/main/SupplementingData_AOOS.pdf), WOD (see https://github.com/SimplySav101/Seascape-Genomics-of-North-Pacific-Forage-Fishes/blob/main/SupplementingData_WODb.pdf), and WorldClim (https://github.com/SimplySav101/Seascape-Genomics-of-North-Pacific-Forage-Fishes/blob/main/SupplementingData_WorldClim.pdf). These guides provide step-by-step instructions for accessing the

databases and scraping the desired parameters (temperature, salinity, pH, chlorophyll, etc.). Generally, we suggest 1) identifying a population genomics dataset for the species and region of interest; 2) parsing collection coordinates and dates; and 3) querying environmental databases for these collection coordinates and dates to return environmental parameters of interest. While broader search parameters can be used to query environmental databases (e.g., collection locality rather than geographic coordinates and/or season rather than collection month and year), more precise collection information will facilitate more accurate environmental measures.

The ‘gold standard’ dataset for seascape genomics is achieved through synchronous environmental data collection with biological sample collection. When this is not possible and a data supplementation approach is taken, reliable collection locations and times must be available. However, even exact collection coordinates and dates will not yield exact matches in environmental databases: every environmental data point supplemented in this way will be a source of error (with closer matches, temporally and spatially, resulting in smaller error). Error will also stem from the requisite trade-off between spatial and temporal proximity of an environmental measurement. Is an environmental measurement more accurate if it is spatially very close to the query coordinates, but temporally distant? What if this temporal distance is interannual (May 2010 vs May 2021) or intrannual (May 2021 vs June 2021)? The answers to these questions will be heavily influenced by the geographic region, environmental parameter(s), and overall objectives of the researcher. Both of these error types – inexactness in query matches and spatiotemporal proximity trade-offs – can be mitigated by avoiding point estimates and instead building a distribution of an environmental measurement at a given set of coordinates over, for example, the past ten years. From this distribution, a summary statistic (mean or median) or set of inferences (n random draws from the distribution) can be used to

describe the environmental conditions at the location. Additionally, forecasting trends is possible using temporal interpolation methods (Lepot et al., 2017) and spatial interpolation models can predict environmental conditions in unsampled areas (reviewed in Li and Heap, 2014).

In summary, while data supplementation is a potentially useful means of leveraging publicly available datasets, it is predominantly a stopgap measure until targeted seascape genomics studies are undertaken. Prior to undertaking such a *post hoc* study, we recommend contacting the authors who generated the population genomics datasets to better understand what lines of inquiry they have pursued, but perhaps not published, and invite their collaboration.

5 Discussion

Seascape genomics is a powerful means of identifying and quantifying the selective pressures exerted on species in a changing climate, providing insight into adaptive potential. Future ecosystem productivity in the North Pacific will be heavily influenced by the continued viability of forage fishes in the region. Management-motivated population genomic studies of North Pacific forage fishes are on the rise (Spies and Punt, 2015; Benestan, 2020), however only one has explicitly tested GEAs (Petrone et al., 2021). In light of the imminent threats of oceanographic shifts caused by climate change and the ecological and economic importance of forage fishes, it is critical that we quantify these associations.

From a conservation perspective, this endeavor becomes even more important as seafood harvested by the Alaska fishing industry represents two thirds of wild seafood harvested by the United States (Alaska Seafood Marketing Institute (ASMI), 2022). The implications of including seascape genomics in fisheries management can diminish the concern for having access to a sufficient quantity of food (Bernatchez et al., 2017): current harvest rates, coupled with the drastic effects of climate change, indicate an unfavorable outlook for forage fishes if comprehensive management techniques are not considered.

We find that substantial resources have already been invested in generating genomic data, and monitoring environmental parameters could empower *post hoc* seascape genomics studies: currently, the NOAA AFSC and the Alaska Department of Fish and Game are transitioning from microsatellite data to genomic datasets (W. Larson, NOAA AFSC; S. Gilk-Baumer, ADF&G; pers. comm.), which will further enable a data supplementation approach. However, as discussed, the supplementation approach we propose serves primarily as a means of beginning to understand where more targeted GEA studies are needed. In the absence of explicit seascape genomics studies, population genomics studies can preserve the

potential for *post hoc* work by consistently providing precise collection dates, coordinates, and depths. An increased focus on synchronous environmental data collection, especially temperature, would constitute an even greater improvement.

Author contributions

LT, SL, JL, and JG contributed to conceptualization, funding requisition/acquisition, and methodology. LT, NT, AR, and SL contributed to data curation and formal analysis. LT, NT, AR, SL, JL, KB, and JG contributed to writing, and editing. LT and NT contributed to visualization and final submission. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by a grant from the NSF Research Coordination Network Evolution in Changing Seas NSF-OCE 1764316. This publication is partially funded by the Cooperative Institute for Climate, Ocean, and Ecosystem Studies (CICOES) under NOAA Cooperative Agreement NA20OAR4320271, contribution no. 2023-1313.

Acknowledgments

We would like to acknowledge our appreciation for the NSF Research Coordination Network Evolution in Changing Seas and the Cooperative Institute for Climate, Ocean, and Ecosystem Studies (CICOES) for their funding support. This is publication #1610 from the Institute of Environment at Florida International University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- 50 CFR Part 679 Authority: 16 U.S.C. 773 et seq.; 1801 et seq.; 3631 et seq.; Pub. L. 108-447; Pub. L. 111-281 (Accessed 20 December 2022).
- Alaska Seafood Marketing Institute (ASMI) (2022) *The economic value of alaska's seafood industry*. Available at: https://www.alaskaseafood.org/wp-content/uploads/MRG_ASMI-Economic-Impacts-Report_final.pdf (Accessed 31 March 2022).
- Benestan, L. (2020). "Population genomics applied to fishery management and conservation," in *Population genomics: marine organisms*. Eds. M. F. Oleksiak and O. P. Rajora (Cham: Springer International Publishing), 399–421.
- Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M., Beacham, T. D., et al. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends Ecol. Evol.* 32 (9), 665–680. doi: 10.1016/j.tree.2017.06.010
- Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., et al. (2016). *NCEI standard product: world ocean database (WOD)* (NOAA National Centers for Environmental Information). Available at: <https://www.ncei.noaa.gov/archive/accession/NCEI-WOD>.
- Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37 (12), 4302–4315. doi: 10.1002/joc.5086
- Galindo, H. M., Olson, D. B., and Palumbi, S. R. (2006). Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Curr. Biol.* 16 (16), 1622–1626. doi: 10.1016/j.cub.2006.06.052
- Gobler, C. J., Merlo, L. R., Morrell, B. K., and Griffith, A. W. (2018). Temperature, acidification, and food supply interact to negatively affect the growth and survival of the forage fish, *Menidia beryllina* (Inland Silverside), and *Cyprinodon variegatus* (Sheepshead Minnow). *Front. Mar. Sci.* 5, 86. doi: 10.3389/fmars.2018.00086
- Hinga, K. R. (1985). Evidence for a higher average primary productivity in the Pacific than in the Atlantic Ocean. *Deep. Sea. Res. Part A. Oceanographic. Res. Papers.* 32 (2), 117–126. doi: 10.1016/0198-0149(85)90023-8
- Hollowed, A. B., Barbeaux, S. J., Cokelet, E. D., Farley, E., Kotwicki, S., Ressler, P. H., et al. (2012). Effects of climate variations on pelagic ocean habitats and their role in structuring forage fish distributions in the Bering Sea. *Deep. Sea. Res. Part II: Topical. Stud. Oceanogr.* 65, 230–250. doi: 10.1016/j.dsr2.2012.02.008
- Lepot, M., Aubin, J. B., and Clemens, F. H. (2017). Interpolation in time series: An introductory overview of existing methods, their performance criteria and uncertainty assessment. *Water* 9 (10), 796. doi: 10.3390/w9100796
- Li, J., and Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Model. Software.* 53, 173–189. doi: 10.1016/j.envsoft.2013.12.008
- Liggins, L., Trembl, E. A., and Riginos, C. (2013). Taking the plunge: an introduction to undertaking seascape genetic studies and using biophysical models. *Geogr. Compass.* 7 (3), 173–196. doi: 10.1111/gec3.12031
- National Marine Ecosystem Status (2020) *Forage fishes and small pelagics*. Available at: <https://ecowatch.noaa.gov/thematic/forage-fish-and-small-pelagics> (Accessed 25 March 2022).
- Orlova, S. Y., Rastorguev, S., Bagno, T., Kurnosov, D., and Nedoluzhko, A. (2021). Genetic structure of marine and lake forms of Pacific herring *Clupea pallasii*. *PeerJ* 9, e12444. doi: 10.7717/peerj.12444
- Ormseth, O. A., and Yasumiishi, E. (2019). Status of forage species in the Bering Sea and Aleutian Islands region. *Alaska. Fisheries. Sci. Center.*, 1–46.
- Petrik, C. M., Stock, C. A., Andersen, K. H., van Denderen, P. D., and Watson, J. R. (2020). Large pelagic fish are most sensitive to climate change despite pelagification of ocean food webs. *Front. Mar. Sci.* 7, 588482. doi: 10.3389/fmars.2020.588482
- Petrou, E. L., Fuentes-Pardo, A. P., Rogers, L. A., Orobko, M., Tarpey, C., Jiménez-Hidalgo, I., et al. (2021). Functional genetic diversity in an exploited marine species and its relevance to fisheries management. *Proc. R. Soc. B.* 288 (1945), 20202398. doi: 10.1098/rspb.2020.2398
- Quintela, M., Bhat, S., Præbel, K., Gordeeva, N., Seljestad, G. W., Hanebrekke, T., et al. (2021). Distinct genetic clustering in the weakly differentiated polar cod, *Boreogadus saida* Lepechin 1774 from East Siberian Sea to Svalbard. *Polar. Biol.* 44 (8), 1711–1724. doi: 10.1007/s00300-021-02911-7
- Selkoe, K. A., Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., et al. (2016). A decade of seascape genetics: contributions to basic and applied marine connectivity. *Mar. Ecol. Prog. Ser.* 554, 1–19. doi: 10.3354/meps11792
- Spies, I., and Punt, A. E. (2015). The utility of genetics in marine fisheries management: a simulation study based on Pacific cod off Alaska. *Can. J. Fisheries. Aquat. Sci.* 72 (9), 1415–1432. doi: 10.1139/cjfas-2014-0050
- Zhang, B. D., Li, Y. L., Xue, D. X., and Liu, J. X. (2020). Population genomics reveals shallow genetic structure in a connected and ecologically important fish from the Northwestern Pacific Ocean. *Front. Mar. Sci.* 7 (2020), 374. doi: 10.3389/fmars.2020.00374