https://doi.org/10.25923/jpth-ec87

NOAA Technical Memorandum NMFS

APRIL 2026

EVALUATION OF TEMPORAL CORRELATION BETWEEN GENOMICS AND POTENTIAL HABITAT PREDICTIONS IN PACIFIC SARDINE (SARDINOPS SAGAX)

Matthew T. Craig[1] and Gary C. Longo[2]

[1] NOAA Fisheries, Southwest Fisheries Science Center Fisheries Resources Division, La Jolla, California

[2] Ocean Associates, Inc. Under Contract to NOAA Fisheries, Southwest Fisheries Science Center Fisheries Resources Division, La Jolla, California

NOAA-TM-NMFS-SWFSC-736

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Southwest Fisheries Science Center

About the NOAA Technical Memorandum series

The National Oceanic and Atmospheric Administration (NOAA), organized in 1970, has evolved into an agency which establishes national policies and manages and conserves our oceanic, coastal, and atmospheric resources. An organizational element within NOAA, the Office of Fisheries is responsible for fisheries policy and the direction of the National Marine Fisheries Service (NMFS).

In addition to its formal publications, the NMFS uses the NOAA Technical Memorandum series to issue informal scientific and technical publications when complete formal review and editorial processing are not appropriate or feasible. Documents within this series, however, reflect sound professional work and may be referenced in the formal scientific and technical literature.

SWFSC Technical Memorandums are available online at the following websites:

SWFSC: https://swfsc-publications.fisheries.noaa.gov/

NOAA Repository: https://repository.library.noaa.gov/

Accessibility information

NOAA Fisheries Southwest Fisheries Science Center (SWFSC) is committed to making our publications and supporting electronic documents accessible to individuals of all abilities. The complexity of some of SWFSC's publications, information, data, and products may make access difficult for some. If you encounter material in this document that you cannot access or use, please contact us so that we may assist you.
Phone: 858-546-7000

Recommended citation

Craig, Matthew T., and Gary C. Longo. 2026. Evaluation of Temporal Correlation between Genomics and Potential Habitat Predictions in Pacific Sardine (Sardinops sagax). U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-736 https://doi.org/10.25923/jpth-ec87

Table of Contents

Introduction

Pacific sardine (*Sardinops sagax*) are a component of the coastal pelagic fish species assemblage in the eastern North Pacific. Historically, Pacific sardine were a major contributor to the U.S. fishing industry and in the 1920s and 1930s the fishery for Pacific sardine was the largest in the Western Hemisphere (Norton and Mason, 2005). Like many other coastal pelagic fish species, Pacific sardine population size fluctuates greatly over both short and long-time scales. These fluctuations seem to be environmentally regulated (Mantua and Hare, 2002), and sardines appear equipped to remain at low population numbers without genetic bottlenecking (Lecomte et al. 2004; Longo et al., 2025).

Understanding spatial population structure is a precondition for sustainable management (Cadrin, 2020). Currently, Pacific sardine in the U.S. are managed as distinct subpopulations that are presumed to be spatially defined, but with overlapping geographic ranges. The Northern Subpopulation (NSP) is thought to range from Alaska, U.S., to northern Baja California, MX. The Southern Subpopulation (SSP) is thought to range from southern California, U.S., to central Baja California, MX (PFMC, 2024). The area of overlap, however, is not thought to be occupied by both groups at the same time; which group occupies the area of overlap depends on environmental conditions. The Pacific Fishery Management Council manages the NSP and attribution of sardine catches and biomass estimates to the NSP in US waters for stock assessment purposes is based upon the sardine potential habitat model (Demer and Zwolinski, 2014; Zwolinski and Demer, 2023). Applying the habitat model for stock assessment purposes has resulted in most sardine in southern California assumed not to be part of the NSP in recent years at a time when the NSP remains at low biomass. Any biomass and catches not attributed to NSP based on the habitat model are excluded from the stock assessment of the NSP (Kuriyama, et al., 2020) upon which management decisions are based.

Using traditional genetic analyses of mitochondrial sequence, nuclear sequence, and nuclear microsatellite data, previous studies have not shown evidence for population structure in Pacific sardine (e.g., Hedgecock, et al., 1989; Lecomte, et al., 2004; Gutierrez-Flores, 2007; Adams and Craig, 2024). These types of data and analyses, however, often have low assignment power which may hinder the detection of genetic partitioning in species with large effective and census population sizes. Recent advances in data generation and analysis (such as the use of single nucleotide polymorphisms and outlier analysis) have greatly improved our ability to identify genetic partitions and genes that are potentially under selection due to environmental conditions (e.g., Grewe, et al., 2015, Vaux, et al., 2020, Teske, et al., 2021). Longo et al. (2025) provided genome-wide analysis of > 9 million SNPs and also found no evidence for genetic structure, evidence of regional selection, or outlier loci in Pacific sardine and concluded that Pacific sardine represents a panmictic population from the Pacific Northwest of the U.S. to the Gulf of California, MX. Following this research finding, the Pacific Fishery Management Council (PFMC) has begun evaluating changes to the Fishery Management Plan (FMP) for the species.

While Longo et al. (2025) found no genetic structure, their samples were all collected during the northern hemisphere summer. Given the hypothesis that the NSP and SSP occupy the waters off

of southern California, USA, at different times of the year (Zwonlinski and Demer, 2023 and references therein), this could be a source of bias. Genome-scale genetic data are able to detect small differences with high accuracy and are thus an ideal tool to use to test the hypothesis that a distinct subpopulation moves in and out of southern California during certain times of year. In this study, we collected sardine each month for a full calendar year and evaluated correlation between possible genetic differences and habitat assignment.

Methods

Sardine were collected through a collaborative research effort between the Southwest Fisheries Science Center and California Wetfish Producers Association. Sampling took place in southern California between the Port of Los Angeles breakwater and Santa Catalina Island (Figure 1, Table 1) on a monthly basis from June, 2022, to May, 2023. Fish were collected from a single purse seine set and 100 individuals were haphazardly selected from the total catch, frozen, and transported to the laboratory. A small piece (~2 x 2 mm) of muscle tissue was removed from the caudal peduncle and stored in 100% ethanol at room temperature until processing.

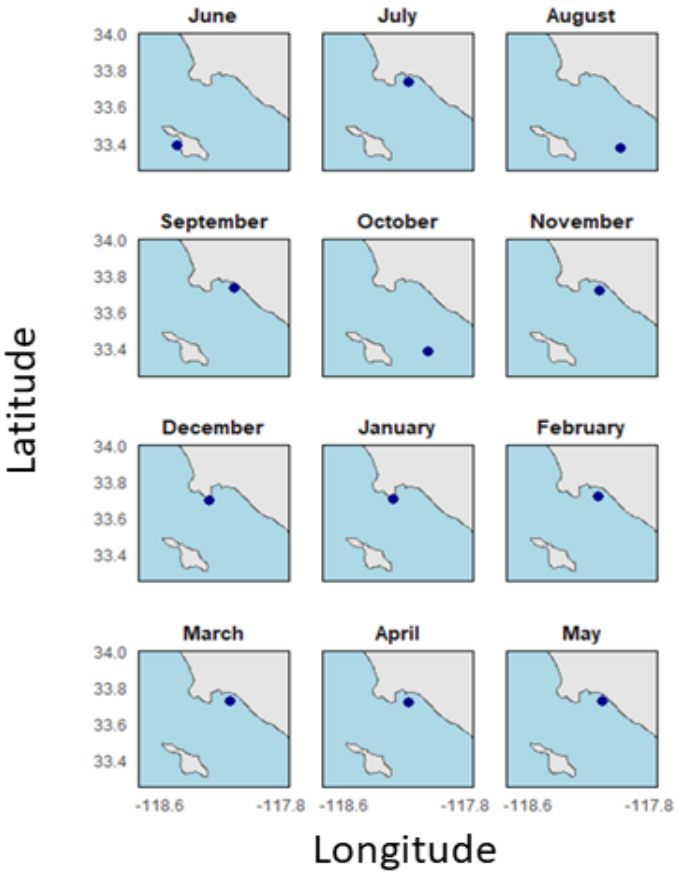

Figure 1. Sampling locations for Pacific sardine (*Sardinops sagax*) from June, 2022, to May, 2023.

Table 1. Sampling locations in decimal degrees for Pacific sardine (*Sardinops sagax*) during 2022-2023.

Month	Latitude	Longitude
June	33.38	-118.49
July	33.73	-118.17
August	33.37	-117.99
September	33.73	-118.12
October	33.38	-118.05
November	33.72	-118.13
December	33.69	-118.28
January	33.70	-118.28
February	33.71	-118.14
March	33.72	-118.14
April	33.71	-118.18
May	33.72	-118.12

DNA was extracted from 50 individuals per month using the Qiagen DNeasy Blood and Tissue 96 kit following manufacturer's protocols. Extracted DNA was run on a 1% agarose gel stained with ethidium bromide to assess DNA quality and was quantified using fluorescence assays on a Perkin-Elmer Victor. 10ng of DNA from each high-quality extraction was plated and allowed to dry at room temperature while covered with a micropore plate cover. DNA was fragmented and tagged (tagmented) with a universal Nextera overhang following the Nextera DNA Library Prep Kit (Illumina, Inc., San Diego, CA) using $1/20^{th}$ of recommended reagents. Tagmented libraries were amplified with low-cycle PCR and barcoded using Illumina Nextera dual-indices at concentrations of 5uM. Illumina P5 and P7 sequencing primers were added using another round of low-cycle PCR. Tagmented and indexed samples were then normalized (<=25ng) using 96-well SequelPrep Normalization Plates following manufacturer protocol and then pooled for each plate. Pooled libraries were cleaned using AMPure XP beads (Beckman Coulter, Inc., Brea, CA) and eluted in 20ul of TLE buffer. Libraries were sequenced on two lanes of 2 x 150bp paired-end Illumina NovaSeq XB+ at Azenta Genewiz. Libraries for samples collected in July, 2022, were created and sequenced prior to this study (Longo et al., 2024) therefore batch effects (i.e., differences attributed to library preparation and/or sequencing) were evaluated (Lou & Therkildsen, 2022).

For lcWGS analyses, haplotype 1 (hap 1) of the Pacific sardine reference genome (Longo et al., 2024; BioProject PRJNA1094947) was indexed using BWA v0.7.17 (Li & Durbin, 2009) and Samtools v1.11 faidx (Li et al., 2009) after excluding contigs that were not incorporated into putative chromosomes. Raw lcWGS data were de-multiplexed into forward and reverse fastq files for each individual. FastQC v0.11.9 (Andrews, 2010) and MultiQC v1.14 (Ewels et al., 2016) was used to check sequence quality of raw reads. Adapters and polyG tails were trimmed from raw fastq files using Trimmomatic v0.39 (Bolger et al., 2014) and fastp v0.23.2 (Chen et al., 2018), respectively. Trimmed reads were aligned to the reference genome using BWA. Samtools was then used to clean up read pairings and flags from BWA with fixmate, convert

sam to bam files, filter non-unique and poor-quality mappings before sorting read pairs by mapping coordinate. After bam files were built, duplicate reads were detected and removed with Picard MarkDuplicates v2.23.9 (http://broadinstitute.github.io/picard/) and overlapping paired-end reads were clipped with bamtools clipOverlap v2.5.1 (Barnett et al., 2011) to generate final bam files. Samtools depth was used to tally alignment depth in all individuals. Individuals with <1x mean depth of coverage were filtered from downstream analyses. To reduce potential sequencing depth bias, we performed targeted down-sampling based on the sequencing library. Target down-sampling depths were drawn from the distribution of mean individual depths calculated from the data.

lcWGS data were initially screened for the presence of Japanese sardine (*Sardinops melanosticta*), a non-native species that was first detected in the California Current Ecosystem in 2022 (Longo et al., 2024). To explore potential genetic structure in our data after removal of any Japanese sardine samples, principal component analysis (PCA) was conducted using PCAngsd (Meisner & Albrechtsen, 2018) using SNPs from the full genome. The covariance matrices were then imported into R (R Core Team, 2024) to perform eigen decomposition and visualization. Previous work indicated that putative chromosomal structural differences were driving observed patterns in the whole genome PCA (Longo et al. 2024, 2025). In order to look for potential population structure outside of structural variation, chromosomes that appeared to harbor chromosomal inversions were removed and PCA was re-run.

Population-level F_{ST} was estimated between two groups representing the putative Northern and Southern subpopulations of Pacific sardine (NSP and SSP, respectively). Samples were assigned to the NSP using the northern subpopulation potential habitat model (Zwonlinski and Demer, 2023). In order to determine weighted pairwise F_{ST} amongst groups, site allele frequency likelihoods were calculated in ANGSD. Global (i.e., between the two subpopulations) and pairwise (among months) F_{ST} were calculated using the folded site frequency spectrum (-realSFS). To assess significance of global F_{ST}, we tested if the observed F_{ST} value fell significantly outside a distribution from permuting individuals, assuming F_{ST} values follow an exponential distribution (Elhaik, 2012).

Results

Initial screening for Japanese sardine identified 8 individuals (four individuals from December, 2022, one from March, 2023, one from April, 2023, and two from May, 2023[1]. After quality filtering and removal of Japanese sardine, 182 individuals remained (Table 2) with a mean coverage of 2.9 (range 1.04 to 5.28). After targeted downsampling, mean coverage was 2.37 (range 1.04 to 3.96). SNP filtering parameters resulted in 10,401,048 polymorphic loci. When seven chromosomes with putative inversions were removed, 7,309,876 polymorphic sites remained (see Longo et al. 2025 for details on putative inversions). Raw data were deposited in GenBank (Biosample PRJNA1094947).

Table 2. Date of collection, subpopulation assignment as determined by the Pacific sardine northern subpopulation potential habitat model, and final sample size for Pacific sardine samples used in this study. NSP = northern subpopulation, SSP = southern subpopulation.

mm_yyyy	subpop	n
06_2022	SSP	11
07_2022	SSP	18
08_2022	SSP	17
09_2022	SSP	17
10_2022	SSP	17
11_2022	SSP	17
12_2022	NSP	10
01_2023	NSP	18
02_2023	NSP	15
03_2023	NSP	16
04_2023	NSP	12
05_2023	NSP	14

Prior to removing chromosomes with putative inversions, PC1 explained 1.7% of the variation in the genome-wide PCA while PC2 explained 0.61% of the variation and individuals were separated into several groups (Figure 2). The PCA groups showed no apparent association with subpopulation assignment or sampling dates.

After chromosomes with putative inversions (see Longo et al. 2025 for details) were removed, no clustering of samples was observed although sample data from July, 2022, occupied a broader PCA space (Figure 3). These individuals were sequenced on a different date therefore this is likely attributed to a batch effect (i.e., artifacts of the DNA sequencing process).

[1] Previously, Japanese sardine were not detected in southern California until the summer of 2023. The detection of Japanese sardine in December, 2022, in the current data represents the earliest detection of Japanese sardine in this region.

The global weighted F_{ST} between putative subpopulations of Pacific sardine was 0.002 and non-significant (P = 0.94). Pairwise comparisons among months were also small and non-significant for all comparisons (Table 3).

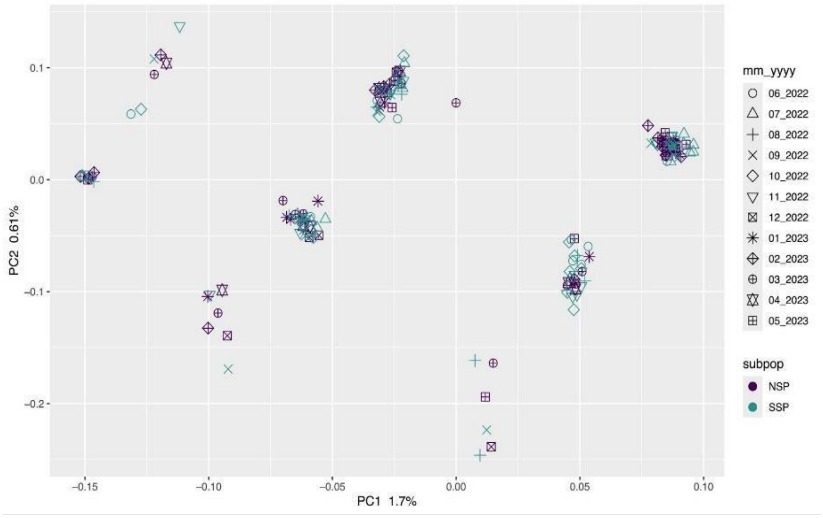

Figure 2. Genome-wide PCA of 10,401,048 polymorphic loci from 182 Pacific sardine, partitioned by sampling month.

Figure 3. Genome-wide PCA of 7,309,876 polymorphic sites of 182 Pacific sardine, partitioned by sampling month, following removal of data from chromosomes with putative inversions.

Table 3. Pairwise comparisons among monthly collections of Pacific sardine (*Sardinops sagax*) collected from June 2022 through May 2023. F_{ST} below diagonal and associated p-values above diagonal.

	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May
Jun.	-	1.000	1.000	1.000	1.000	0.823	1.000	1.000	1.000	1.000	1.000	1.000
Jul.	0.003	-	0.723	0.804	0.827	0.746	0.780	0.936	0.907	0.806	1.000	0.944
Aug.	0.003	0.006	-	0.860	0.921	0.595	0.523	0.963	0.827	0.839	0.812	0.712
Sep.	0.003	0.006	0.005	-	0.921	0.738	0.554	1.000	0.954	0.974	0.906	0.720
Oct.	0.003	0.005	0.005	0.005	-	0.776	0.661	1.000	0.941	1.000	0.981	0.886
Nov	0.004	0.006	0.007	0.006	0.005	-	0.490	0.923	0.689	0.795	0.680	0.766
Dec.	-0.001	0.006	0.006	0.006	0.005	0.006	-	0.923	0.391	0.515	0.000	0.198
Jan.	0.002	0.005	0.005	0.005	0.004	0.005	0.005	-	1.000	1.000	1.000	1.000
Feb.	0.001	0.005	0.005	0.005	0.004	0.006	0.004	0.004	-	1.000	1.000	0.855
Mar.	0.001	0.006	0.005	0.004	0.004	0.005	0.005	0.004	0.004	-	1.000	0.951
Apr.	-0.002	0.005	0.005	0.005	0.004	0.006	0.002	0.004	0.003	0.003	-	0.772
May	0.000	0.005	0.005	0.005	0.004	0.005	0.004	0.004	0.004	0.004	0.003	-

Discussion

In addition to the results presented by Longo et al. (2025), the absence of genetic differences in the genomic architecture of sardine collected on a monthly basis for one year further demonstrates that sardine assigned to the NSP do not represent a genetically distinct unit. While this result is not unexpected given that panmixia from Washington state, USA, to the Gulf of California, Mexico, was demonstrated by Longo et al. (2025), it confirms that those results were not biased by using samples collected primarily during the northern hemisphere summer. These results further support the notion that the designation of sardine sub-populations as fishery management units represents a misalignment of genetic and management units. Such misalignment, or incoherent dimensionality (Berger et al., 2021), in fisheries management can have detrimental effects and may result in inaccurate reference points used for management. The PFMC is currently considering amending the relevant FMP to recognize a single management unit for sardine in U.S. waters (PFMC, 2025). The adoption of that change would be consistent with the results reported here and will ensure that sardine management and genetic units are coherent.

Acknowledgements

The authors wish to thank Diane Pleschner-Steele and Mark Fina for coordinating the cooperative research activities for this study and for their steadfast dedication to collaborative research. We also thank captains Jamie and Richie Ashley and the crew of the F/V Provider for their efforts to collect sardine for this study. We would like to thank Wes Larson, John Hyde, and Annie Yau for reviews of this manuscript.

References

Adams, E.S., Craig, M.T. 2024. Phylogeography of the Pacific sardine, *Sardinops sagax*, across its northeastern Pacific range. Bull. Southern California Acad. Sci. 123(1):10-24.

Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., Marth, G.T., 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27, 1691–1692. https://doi.org/10.1093/bioinformatics/btr174

Berger, A. M., J. J. Deroba, K. M. Bosley, et al. 2021. IncoherentDimensionality in Fisheries Management: Consequences of MisalignedStock Assessment and Population Boundaries. ICES Journal of Marine Science 78: 155–171. https:// doi.org/10.1093/icesjms/fsaa203.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Cadrin, S. X. 2020. Defining spatial structure for fishery stock assessment. Fisheries Research, 221: 105397.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Demer, D.A., and Zwolinski, J.P. 2014. Corroboration and refinement of a method for differentiating landings from two stocks of Pacific sardine (*Sardinops sagax*) in the California Current. ICES Journal of Marine Science 71(2): 328–335. doi: 10.1093/icesjms/fst135.

Elhaik, E., 2012. Empirical Distributions of FST from Large-Scale Human Polymorphism Data. PLoS ONE 7, e49837. https://doi.org/10.1371/journal.pone.0049837

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Grewe, P., et al. 215. Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. Scientific Reports, 5:16916.

Gutiérrez-Flores, C. 2007. Estructura genética poblacional de la sardina del pacífico nororiental *Sardinops sagax caeruleus*, Tesis de Maestría, CICESE, Ensenada, B.C, 2007, 112pp.

Hedgecock, D., et al. 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*, Fishery Bulletin 87(3) (1989), 653-671.

Kuriyama, P. T., J. P. Zwolinski, K. T. Hill, and P. R. Crone. 2020. Assessment of the Pacific sardine resource in 2020 for U.S. management in 2020-2021. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-628.

Lecomte, F., et al. 2004.Living with uncertainty: genetic imprints of climate shifts in East Pacific anchovy (*Engraulis mordax*) and sardine (*Sardinops sagax*). Molecular Ecology, 13(8), 2169-2182.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Longo, G., Minich, J., Allsing, N., James, K., Adams-Herrmann, E., Larson, W., Hartwick, N., Duong, T., Muhling, B., Michael, T. and Craig, M. (2024), Crossing the Pacific: Genomics Reveals the Presence of Japanese Sardine (*Sardinops melanosticta*) in the California Current Large Marine Ecosystem. Mol Ecol, 33: e17561. https://doi.org/10.1111/mec.17561

Longo, G. C., K. D′Amelio, W. Larson, et al. 2025. "Population Genomics Reveals Panmixia in Pacific Sardine (*Sardinops sagax*) of the North Pacific." Evolutionary Applications 18, no. 9: e70154. https://doi.org/10.1111/eva.70154.

Mantua, N. J., & Hare, S. 2002. Pacific-Decadal Oscillation (PDO). Encyclopedia of global environmental change, 1, 592-594.

Meisner, J., and A. Albrechtsen. 2018. "Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data." Genetics 210, no. 2: 719–731.

Norton, J. G., & Mason, J. E. 2005. Relationship of California sardine (*Sardinops sagax*) abundance to climate-scale ecological changes in the California Current system. California Cooperative Oceanic Fisheries Investigations Report, 46, 83.

PFMC, 2024. Coastal Pelagic Species Fishery Management Plan. https://www.pcouncil.org/fishery-management-plan-and-amendments/

PFMC, 2025. Council Summary Decision Document, November, 2025. https://www.pcouncil.org/november-2025-decision-summary-document

R Core Team. 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Teske, P. R., et al. 2001. The sardine run in southeastern Africa is a mass migration into an ecological trap. Science Advances 2021; 7:eabf4514.

Vaux, F., et al. 2021. Adaptive markers distinguish North and South Pacific Albacore amid low population differentiation. Evolutionary Applications 2021; 00:1-22.

Zwolinski, J. P., and D. A. Demer. 2023. An updated model of potential habitat for northern stock Pacific Sardine (*Sardinops sagax*) and its use for attributing survey observations and fishery landings. Fisheries Oceanography, 33:1-14.