Environmental Studies Program: Ongoing Study

Title	Demonstration Project, Integrating DNA Profiles, Genomics and Photo-Identification data in long term monitoring of long lived marine megafauna (NT-17-03)
Administered by	Headquarters
BOEM Contact(s)	Jacob Levenson (jacob.levenson@boem.gov)
Procurement Type(s)	Cooperative Agreement
Conducting Organization(s)	Oregon State University
Total BOEM Cost	$350,000
Performance Period	FY 2020–2022
Final Report Due	September, 2022
Date Revised	August 27, 2021
PICOC Summary	
Problem	A growing number of large-scale, long-term studies of marine mammals and other marine megafauna (e.g., sharks, and turtles) are collecting spatially explicit records linked through individual identification to genetic samples, photo-identification and telemetry. Individual-based studies of long-lived species, require standardization of protocols and distributed access to integrated databases of relevant metadata from multiple studies, including DNA profiles, genomic profiles, photo-identification records and satellite telemetry tracks. For BOEM, the lack of agreed community standards or a collaborative framework for these datasets reduces the potential for long-term monitoring of the health and life history parameters of populations exposed to chronic or acute anthropogenic impacts of BOEM permitted activities.
Intervention	Would standardization and integration of the genetic & photoID allow for an improved sex-specific estimate of abundance.
Comparison	The predicted outcome could be improved identification of population segments and estimates of abundance, especially among largely isolated populations such as the Gulf of Mexico Sperm Whale.
Outcome	The predicted outcome could be improved identification of population segments and estimates of abundance, especially among largely isolated populations such as the Gulf of Mexico Sperm Whale.
Context	Nationwide coastal and offshore waters, however a pilot project could take place among any region, for example, Gulf of Mexico Sperm Whales.

BOEM Information Need(s): Numerous types of studies have been conducted on cetaceans in the U.S. outer continental shelf (OCS), leading to challenges in the analysis and synthesis of diverse datasets, particularly those from long-term studies of populations exposed to anthropogenic impacts. The Marine Mammal Protection Act, National Environmental Policy Act and the Endangered Species Act require BOEM to consider the impacts of energy and mineral extraction activities on cetaceans. Together with other agencies, BOEM invested considerably in studies where a number of groups have collected biopsy

samples for genetic analyses and photographs for individual identification. In addition, many cases of OCS research across program areas and regions included satellite telemetry for studies of habitat use. For example, over the last two decades, several projects have collected biopsy samples and photo-identification records from sperm whales during tagging deployment (e.g., the Sperm Whale Seismic Study or SWSS in 2002) and similar projects are ongoing. DNA profiles and photo-identification records are increasingly used for defining units of management and for estimating abundance and trends in populations using capture-recapture models. To date, however, there is no single catalogue or accessible archive that integrates DNA profiles and photo-identification for individual identity of most of these populations. A Web-based, distributed-access data archive could consolidate and integrate these data greatly enhancing the potential for long-term monitoring. This type of effort will enable and guide new analyses to inform future management questions regarding cetacean populations, including abundance, range, distribution, and stock identity. A similar information need has been recognized by the Office of Naval Research in funding development of the *geneGIS* tools for the SPLASH program on humpback whales in the North Pacific. However, no one is currently uniting the different groups working on including genetics in population assessments such as, for example, GOM Sperm Whales.

Decision Context: In a draft biological opinion as part of an interagency consultation, NMFS does not refer to the GOM sperm whales as a distinct population segment. This likely is a flawed evaluation of the distinctiveness of the Gulf of Mexico population and fails to recognize the uniqueness of both acoustics and mating system of sperm whales (Alexander, et al 2016) The BiOp additionally could have leveraged information from a wealth of Sperm whale studies happening in the regions as well as how genotyping is readily available and could be used in environmental assessments, however, to do so, requires standardized database architecture**.** Lastly, the draft BiOp emphasized the need for improved monitoring of long-term population status and highlights how unsatisfactory the current evidence is.

Information needs met by this study include standardization and archiving of DNA profiles from past and ongoing studies, e.g., mtDNA haplotypes for maternal lineages, sex identification and microsatellite genotypes for individual identification. The DNA profiles derived from multi-locus[1] genetic markers are typically laboratory specific and their format is not compatible with public genetic archives such as GenBank. This project will increase management effectiveness and improve estimates of abundance based on either photo-identification or genotype capture-recapture, complementing the conventional use of genetic markers for describing stock structure. Life history parameters based on mortality can also be derived from capture-recapture models. Visual exploration and spatial analyses of individual-based records from DNA profiles and photo-identification records will be improved. This enables investigators and managers to better understand the population structure and abundance as well as mortality events.

This proposed project ties-in closely with existing efforts to monitor cetacean ecology in the GOM, Alaska and Atlantic regions, by creating a database framework that affords a directory to biological sample analysis across varying studies using a common Web framework. Information obtained will ensure that BOEM complies fully with environmental regulation and considers the impacts of its programs, on these endangered species while building capacity for further study by other BOEM regions using the same infrastructure.

A growing number of large-scale studies of marine mammals and other marine megafauna (e.g., sharks, and turtles) are collecting spatially explicit records linked through individual identification to genetic samples, photo-identification and telemetry. These spatio-temporal records have been used to track the migration and life history parameters of individuals, to estimate the abundance and trends of

populations and, in the case of genetic markers, to infer close kinship (e.g., parent/offspring relationships) and define management units, or Distinct Population Segments. The skin and blubber biopsy samples collected for genetic analyses are also used for assaying a growing number of ecological markers (e.g., stable isotopes, fatty acids) and environmental contaminants (e.g., Persistent Organic Pollutants, heavy metals). To date, however, there has been a conspicuous absence of integration and spatial exploration of individual genetic and photo-identification records; in particular, there is a need for linking photo-identification to genetic information (e.g., DNA profiles).

An integrated database of photo-identification and DNA profiles can be used to manage OCS resources by, a) aiding in long-term monitoring of population health, b) enabling easy, open access to data for conducting analysis to support environmental impact statements and c) accounting for more accurate geographic analysis of DNA biopsy samples. This is essential for monitoring habitat use, assessing environmental threats, and defining appropriate listing levels for species exposed to exploration and development. This benefits BOEM by providing the potential for improved estimates of abundance and mortality/survival based on capture/recapture analyses using existing photo-identification and DNA profiles. This, along with improved genetic information on stock identity, can be used to estimate population losses. Additionally, this will develop a data archiving model that can be adopted nationally. The software architecture of the online archive could also provide tools for primary analyses and visualization of spatial-temporal records, linked through individual identity and genetic markers, as well as export functions for more specialized programs. populations.

Web-based databases are the solution for extending the range of analyses in space and time, standardizing data formats to ensure long-term archive utility, offering a cost-effective data-management solution for individual researchers that robustly provides the analysis functions they require, and encourages collaboration among diverse groups of investigators, thus helping establish community standards for these ongoing studies of long-lived species.

Objectives: The objective of this study is to improve the current computation capabilities for integrating DNA profiles with photo-identification records for assessment and long-term monitoring of marine mammal populations. To provide standardized database architecture for online access, with associated tools for primary analyses and visualization of spatial-temporal records. An important component of this objective is to develop 'community standards' that can be applied across multiple species, as well as multiple investigators, and that promotes data sharing and archiving. The development of a Web-based, collaborative collection could form a model for similar developments with other species of interest to BOEM.

Methods: A Web-based, user-friendly computational framework for accessing integrated DNA profiles and photo-identification records, compatible with the OBIS-SEAMAP, will be developed. Associated computational tools for data exploration and export to specialized programs will also be developed. The proposed work will include organizing a workshop of Federal partners, academic researchers and other potential contributors to the databases, to encourage participation in the data archive and promote collaborative research. Data format commonalities and/or conversion needs will also be determined.

Specific Research Question(s): This profile address several subject areas identified in BOEM's ESP Strategic Framework. This includes Affected Resources Information on the status, trends, and resilience of potentially impacted natural and cultural resources...particularly 'highly regulated and vulnerable to adverse change in status'. Specific research questions: Can integration of Genomic samples when coupled with telemetry and mark-recapture data increase accuracy and improve long term monitoring.

Can computational genomics aid BOEM in better using existing or emerging technology to achieve more effective or efficient scientific results?

Current Status: The project was initiated on 22 September, 2020, with a virtual "kickoff" meeting of the BOEM administrative team on 22 October, 2020. A virtual kickoff meeting of the collaborating data providers was held on 8 December, 2020. The COVID 19 pandemic has created a number of challenges for the PI and data providers. Specifically, federal and university facilities are in lock-down or partial lock-down with restricted access to genetic samples and data archives. Despite these restrictions, we have made slow progress through virtual meetings and limited access to laboratory facilities.

Affiliated WWW Sites: None

References:

Dick DM, Walbridge S, Wright DJ, Calambokidis J, Falcone EA, Steel D, Follett T, Holmberg J, Baker CS (2014) geneGIS: Geoanalytical Tools and Arc Marine Customization for Individual-Based Genetic Records. Transactions in GIS 18: 324-350

Alexander, A., D. Steel, K. Hoekzema, S.L. Mesnick, D. Engelhaupt, I. Kerr, R. Payne and C.S. Baker. 2016. What influences the worldwide genetic structure of sperm whales (*Physeter macrocephalus*)? Molecular Ecology 25:2754–2772.