

# NOAA 'Omics White Paper: Informing the NOAA 'Omics Strategy and Implementation Plan

Kelly Goodwin, Rebecca Certner, Mark Strom, Felipe Arzayus, Margot Bohan, Shalin Busch, Gabrielle Canonico, Scott Cross, Jeanette Davis, Katharine Egan, Thomas Grieg, Ed Kearns, Jennifer Koss, Kirsten Larsen, David Layton, Krista Nichols, Jonathan O'Neil, Derek Parks, Leanne Poussard, Cisco Werner

doi: 10.25923/bd7z-zb37

## ABOUT THIS DOCUMENT

This document (“white paper”) was developed by the National Oceanic and Atmospheric Administration (NOAA) 'Omics Task Force (OTF), a cross-NOAA interdisciplinary team established by the NOAA Research Council. The OTF was tasked in 2018 with documenting the portfolio of 'omics research and activities across the organization, identifying priorities, fostering integration and communication, and formulating solutions to implementation challenges in this nascent sector of research. This document informed the creation of the 2020 [NOAA 'Omics Strategy](#), [NOAA 'Omics Fact Sheet](#), and NOAA 'Omics Implementation Plan and thus provides important supplemental information to those documents.



## CONTENTS

List of Abbreviations	4
Executive Summary	7
Overview	9
Introduction	9
Advance Bioinformatics and Analysis Capacity	11
Enhancement of a Well-Trained Workforce	11
Investment in Bioinformatics Infrastructure and Analysis Capacity	12
Targeted and Coordinated Research	13
Partnerships	14
Accelerate ‘Omics to Applications	15
Research Highlight: Environmental DNA (eDNA)	17
‘Omics Broader Impacts	18
Support the Blue Economy	19
Promote Operational Improvement	19
Advance Ocean Science and Stewardship	19
Summary and Next Steps	20
NOAA ‘Omics Strategy Goals and Objectives	21
Research Highlight: Environmental DNA (eDNA)	22
Table 1. Implementation Status of ‘Omics Technology Applications	24
Appendix I. Portfolio Analysis	25
National Marine Fisheries Service (NMFS) Activities in ‘Omics	25
Bioinformatics analysis and needs in NOAA Fisheries Science Centers	27
NMFS Strategic Initiative on ‘Omics (FY19-FY24)	27
Ocean and Atmospheric Research (OAR) Activities in ‘Omics	28
Aligned OAR ‘Omics Laboratories	28
Atlantic Oceanographic & Meteorological Laboratory (AOML) ‘Omics Program	30
Pacific Marine Environmental Laboratory (PMEL) ‘Omics Laboratory	36
Great Lakes Environmental Laboratory (GLERL) ‘Omics Program	37
Ocean Acidification Program (OAP) Investment in ‘Omics	38
Office of Ocean Exploration and Research (OER) ‘Omics Interest and Investment	38
Sea Grant Investment in ‘Omics	39
National Ocean Service (NOS) Activities in ‘Omics	41
‘Omics at the Hollings and Beaufort Marine Laboratories	41
Coral Reef Conservation Program (CRCP) Interest and Investment in ‘Omics	42
United States Integrated Observing System (U.S. IOOS) Investment in ‘Omics	43
United States Marine Biodiversity Observation Network (U.S. MBON)	43
NOAA Information Enterprises and ‘Omics	44
NOAA External Engagement in ‘Omics	45
National Conference on Marine Environmental DNA	45
U.S.-Norway Intergovernmental Group on eDNA	46
Atlantic Ocean Research Alliance (AORA) Marine Microbiome Group	46
Global ‘Omics Observation Network (GLOMICON)	46

NOAA Cross-Agency Engagement in ‘Omics	47
Microbiome Interagency Working Group (MIWG)	47
Interagency Working Group on Biological Data Sharing (IWGBDS)	47
Government eDNA Working Group (GEDWG)	47
Appendix II. ‘Omics Task Force Survey Results and Infrastructure Review	48
Survey Results	48
Review of NOAA Infrastructure to Support ‘Omics	54
Appendix III. Summary of Terms and Glossary	57
Appendix IV. Contributors	59
Appendix V. References	61

## **LIST OF ABBREVIATIONS**

AFSC – Alaska Fisheries Science Center  
AI – Artificial Intelligence  
AOML – Atlantic Oceanographic and Meteorological Laboratory  
AORA – Atlantic Ocean Research Alliance  
AUV – Autonomous Underwater Vehicle  
AWS – Amazon Web Services  
BOEM – Bureau of Ocean Energy Management  
CalCOFI – California Cooperative Oceanic Fisheries Investigations  
CCEHBR – Coastal Environmental Health and Biomolecular Research  
CI – Cooperative Institute  
CKMR – Close-kin Mark Recapture  
CRADA – Cooperative Research and Development Agreement  
CRCP – Coral Reef Conservation Program  
CPO – Climate Program Office  
CSP – cloud service provider  
ddRAD-seq – Double Digest Restriction Associated DNA Sequencing  
DNA – deoxyribonucleic acid  
DoD – Department of Defense  
ECO – Ecosystem and Community ‘Omics  
eDNA – environmental DNA  
EDMC – Environmental Data Management Committee  
EMP – Earth Microbiome Project  
EPP – Educational Partnership Program  
ESA – Endangered Species Act  
ESP – Environmental Sample Processor  
FFO – federal funding opportunity  
FTE – full time employee  
GBIF – Global Biodiversity Information Facility  
GEDWG – Government eDNA Working Group  
GLERL – Great Lakes Environmental Research Laboratory  
GLOMICON – Global ‘Omics Observation Network  
HAB – Harmful Algal Bloom  
HML – Hollings Marine Laboratory  
HMS – highly migratory species  
HPCC – High-Performance Computing Cluster  
IOC – Intergovernmental Oceanographic Commission  
IOOS – Integrated Ocean Observing System  
IT – Information Technology  
IWGBDS – Interagency Working Group on Biological Data Sharing  
IUU – Illegal, Unreported, and Unregulated  
JCVI – J. Craig Venter Institute  
LAUV – long-range autonomous underwater vehicle

LOTMC – Line Office Transition Managers Committee  
MBARI – Monterey Bay Aquarium Research Institute  
MBON – Marine Biodiversity Observation Network  
MIWG – Microbiome Interagency Working Group  
ML – Machine Learning  
MSI – Minority Serving Institution  
MST – Molecular Source Tracking  
NASA – National Aeronautics and Space Administration  
NCCOS – National Centers for Coastal Ocean Science  
NCEI – National Centers for Environmental Information  
NCOG – NOAA CalCOFI Ocean Genomics  
NEFSC – Northeast Fisheries Science Center  
NESDIS – National Environmental Satellite, Data, and Information Service  
NGI – Northern Gulf Institute  
NIH – National Institutes of Health  
NMDC – National Microbiome Data Collaborative  
NMFS – National Marine Fisheries Service  
NMITS – NOAA Mission Information Technology Services  
NMNH – Smithsonian National Museum of Natural History  
NOAA – National Oceanic and Atmospheric Administration  
NOPP – National Oceanographic Partnership Program  
NOS – National Ocean Service  
NRAP – NOAA Rotational Assignment Program  
NRC – NOAA Research Council  
NRDD - NOAA Research and Development Database  
NSF – National Science Foundation  
NSL – National Systematics Laboratory  
NSSL – National Severe Storms Laboratory  
NWFSC – Northwest Fisheries Science Center  
NWS – National Weather Service  
OAP – Ocean Acidification Program  
OAR – Oceanic and Atmospheric Research  
OBIS – Ocean Biodiversity Information System  
OCIO – Office of the Chief Information Officer  
OEC – ‘Omics Executive Committee  
OER – Office of Ocean Exploration and Research  
OMAO – Office of Marine and Aviation Operations  
ONR – Office of Naval Research  
OTF – ‘Omics Task Force  
OWG – ‘Omics Working Group  
PMEL – Pacific Marine Environmental Laboratory  
POGO – Partnership for Observation of the Global Ocean  
qPCR – quantitative polymerase chain reaction  
R&D – Research and Development

RNA – ribonucleic acid  
SAB – Science Advisory Board  
SAS – subsurface automatic samplers  
SBIR – Small Business Innovation Research  
SEFSC – Southeast Fisheries Science Center  
SIO – Scripps Institution of Oceanography  
SNP – single nucleotide polymorphism  
SOST – Subcommittee on Ocean Science and Technology  
STAR – Center for Satellite Applications and Research  
SWFSC – Southwest Fisheries Science Center  
TPO – Technology Partnerships Office  
TSMRI – Ted Stevens Marine Research Institute  
UAV – Unmanned Aerial Vehicle  
USACE – United States Army Corps of Engineers  
USDA – United States Department of Agriculture  
USFWS – United States Fish and Wildlife Service  
USGS – United States Geological Survey  
UxS – Unmanned Systems

# Executive Summary

‘Omics<sup>1</sup> refers to a suite of technologies used to analyze DNA, RNA, proteins, or metabolites. These tools can be used to identify organisms (*who is there*), understand their behavior (*what they are doing*), and predict shifts in their population structure (*how they might adapt or decline in a changing ocean*). Advances in ‘omics offer improvements to the cost, accuracy, breadth, and speed of biological data delivery compared to traditional observational techniques. Investment in these technologies will strengthen NOAA understanding, mapping, and management of economically valuable species, vulnerable populations, and harmful organisms. NOAA recognizes the importance of utilizing ‘omics science to enhance the agency’s monitoring, modeling, and warning systems, all of which are critical to fulfilling our missions and support the Blue Economy<sup>2</sup>.

Effective management of ocean ecosystems and the sustainable use of marine resources requires leveraging new technologies and scientific innovation. NOAA’s diverse missions depend on an equally diverse set of skills, a constantly evolving workforce, and utilization of state-of-the-art technologies. The beneficial range and volume of information generated through ‘omics tools will allow NOAA to comprehensively address a complex suite of questions across multiple missions. Although significant challenges with ‘omics remain, these tools have already transformed our ability to characterize and quantify the biological composition of the ocean.

This document describes a suite of ‘omics efforts in NOAA and recommends three specific areas for investment and development:

1. *Enhancement of a Well-Trained Workforce*: increase NOAA’s capacity and expertise for ‘omics data collection, processing, and transition by providing bioinformatics<sup>3</sup> training for its scientists and IT professionals, and by recruiting ‘omics expertise into the federal workforce.
2. *Investment in Bioinformatics Infrastructure and Analysis Capacity*: provide adequate data storage, accessibility, processing, and analytical capacity to meet NOAA’s requirements for throughput and efficiency.
3. *Prioritization of Targeted and Coordinated Research*: leverage programs and invest in research that can support and benefit from process oriented, as well as routine collection of ‘omics samples.

---

<sup>1</sup> A collective suite of technologies used to investigate biological molecules (DNA, RNA, proteins, metabolites) that translate into the identity, function, and relationship dynamics of an organism(s).

<sup>2</sup> The Blue Economy includes the resources and services the ocean provides and from which society derives economic benefits, such as fisheries, aquaculture, transportation, and tourism. NOAA’s investment in science underpins the Blue Economy’s management and sustainable use of ocean resources in support of national, economic, and environmental security.

<sup>3</sup> Refers to the blending of biology and computer science necessary to process and interpret large ‘omics data sets.

In addition, this white paper suggests a timeline to transition from ‘omics research and technology to use and application, and to identify some of the broader impacts of ‘omics technologies and benefits to economic growth in the Blue Economy. Carrying out the described investment areas requires intra-agency coordination and external cross-collaboration with Federal, private, and academic partners.

This white paper, developed primarily in 2018, was used to inform the 2020 [NOAA ‘Omics Strategy](#), the [NOAA ‘Omics Fact Sheet](#), and the NOAA ‘Omics Implementation Plan. These documents seek to enhance and maximize ‘omics efforts across the agency and to improve our knowledge to better meet NOAA’s mission to explore, define, and characterize the biological components of the ocean, and to manage, protect, and restore living marine resources. The NOAA ‘Omics Strategy developed and refined the ideas provided here and presented the information in a format that was standardized among the [NOAA Science and Technology Focus Areas](#) of [‘Omics](#), [Unmanned Systems](#), [Artificial Intelligence](#), and [Cloud Computing](#). The NOAA ‘Omics Implementation Plan details how the goals and objectives defined in the NOAA ‘Omics Strategy can be implemented and accomplished on an agency-wide scale. The Implementation Plan outlines actionable investments that facilitate the integration of ‘omics into wider NOAA mission areas. The reader is directed to the [NOAA ‘Omics Strategy](#) and [NOAA ‘Omics Fact Sheet](#), which were developed after the creation of this document.



# NOAA ‘Omics White Paper: Informing the NOAA ‘Omics Strategy and Implementation Plan

## *Overview*

A portfolio analysis showed that NOAA is applying ‘omics approaches to the ocean sciences to benefit a variety of missions including: fisheries management, aquaculture development, food and water safety, species and habitat conservation, seafood consumer protection, and biological discovery (Appendix I). In broad summary, current and planned applications of molecular approaches (Table 1) accelerate ecosystem understanding and promote agency efficiency and the Blue Economy. While the portfolio analysis revealed impressive breadth and depth to NOAA’s utilization of cutting edge ‘omics tools, impediments that need to be addressed were also uncovered. Hurdles to achieving desired outcomes include insufficient staffing, internal ‘omics expertise, and infrastructure for data analysis and sequencing (Appendix II). Recommendations to accelerate the state of the science include workforce enhancement, increased analysis capacity and sequencing infrastructure, and support for targeted research carried out in a coherent and coordinated fashion.

## *Introduction*

‘Omics, the suite of technologies used to investigate biological molecules that translate into the identity, function, relationship dynamics of organisms, has rapidly increased in capacity, speed, and sensitivity, while at the same time decreasing in cost. Techniques such as genomics, transcriptomics, proteomics, and metabolomics use DNA, RNA, proteins, and various small molecules to identify organisms and their responses to environmental conditions. Advances in sequencing technology have transformed the biological sciences and fostered new insights into everything from health monitoring to economic markets (e.g., microbiome-based medicines, fuel cells, and biotechnology/industrial applications).

The NOAA Science Advisory Board (SAB) provided a number of recommendations directly relevant to NOAA ‘Omics (see Yoder et al. 2016). The SAB’s “*Adopt Now*” recommendations included: invest in genomic technologies and their applications, develop new applications for environmental DNA (eDNA) monitoring and analysis (see Box inset below), utilize autonomous vehicles and sensor technologies for measurements related to living marine resources, and support sophisticated data analysis techniques and personnel trained to use them. The present document reflects these recommendations, with use cases and approaches provided in the Sections below and summarized in Table 1.

NOAA is already applying ‘omics approaches in support of mission areas including fisheries management, aquaculture, food and water safety, protected species and habitat conservation, seafood consumer protection, and biological discovery. For more information, see NOAA’s response to the SAB’s recommendations (Werner et al. 2018), and Appendices I and II for results of a cross-Line Office portfolio analysis conducted by the ‘Omics Task Force<sup>4</sup> (OTF). The OTF found that molecular approaches currently applied to NOAA’s mission include population genetics, restoration genetics, eDNA, and absolute abundance estimation by close-kin mark-recapture. Applications include estimating catch compositions of mixed-stock fisheries, detection and control of invasive species, protection of critical breeding populations, enumeration of pathogens and harmful algae, detection of microbial/geochemical shifts, recording regional and large-scale changes to biological pathways that affect an array of ecosystem services, and understanding genetic factors that control pathogen virulence with consequences to commercial fishing, aquaculture, and animal health. Furthermore, as ship-time and labor costs increase, ‘omics technologies promise efficiencies because they are amenable for use *in situ* and on autonomous platforms that relay information in near real-time, and do not rely on labor-intensive sorting techniques. These developments are timely because although the rate of change in the global ocean is concerning, no certain baseline exists against which this change can be measured. The emergence of ‘omics can help deliver baseline biological observations.

Building on these activities, the OTF identified opportunities and needs for investments in presently ongoing and anticipated (next 5-10 years) applications of molecular/‘omics approaches (Table 1). The OTF also identified limitations to fully integrating ‘omics into NOAA research and operations. Gaps were noted in staffing, information technology, data analytics, computing infrastructure, and bioinformatics expertise. Recommendations include sustained (multi-year) support to provide needed infrastructure for research and development, and ultimately for successful transition to operations and applications. These investments are viewed as necessary to contribute to accelerating our understanding of marine ecosystems and associated services, and to promote the agency’s effective support of the Blue Economy. This information and an additional series of scoping exercises were used to guide the NOAA ‘Omics Strategy (OSET 2020) and NOAA ‘Omics Implementation Plan (Goodwin et al. 2020). Surveys of programs and line offices detailing current ‘omics activities are outlined in Appendix I for use in the strategy. Additional scoping was conducted via surveys at the line office and program level to further identify current and needed future ‘omics activities for use in the implementation plan.

## **Advance Bioinformatics and Analysis Capacity**

NOAA ‘Omics seeks to enhance and maximize ‘omics efforts that improve our knowledge to better meet NOAA’s mission to manage, protect, and restore living marine resources. Based on the OTF’s analyses, recommended areas of investment for NOAA’s ‘omics research include:

---

<sup>4</sup> The NOAA ‘Omics Task Force (OTF) is a cross-NOAA interdisciplinary team established by the NOAA Research Council in 2018. The OTF was tasked to document the portfolio of ‘omics research and activities across the organization, identify priorities, foster integration and communication, and formulate solutions to implementation challenges in this nascent sector of research. See Appendix 3 for the OTF’s membership.

- Enhancement of a well-trained workforce,
- Investment in bioinformatics infrastructure and data analytics capacity, and
- Prioritization of targeted and coordinated research.

These broad areas have been further developed into the objectives available in the NOAA 'Omics Strategy. The information below provides supplemental information that is consistent with and supports the NOAA 'Omics Strategy and the NOAA 'Omics Implementation Plan. The reader is directed to the [NOAA 'Omics Strategy](#) and [NOAA 'Omics Fact Sheet](#), which were developed after the primary creation of this document.

## **Enhancement of a Well-Trained Workforce**

To address gaps in the ability to analyze and provide environmental context to sequence data, NOAA needs to recruit and retain early-career scientists and information technology (IT) professionals. Note that presently there are no known NOAA FTEs who specialize in bioinformatics. Post-doctoral support can recruit cutting-edge genomic research and bioinformatics expertise from academia. In addition, full-time equivalent positions are required to achieve laboratory and IT enhancements that are operational and able to evolve with advancing technologies. NOAA should collaborate with academic and private sectors throughout the internal capacity building process including the Cooperative Institute network and NOAA education and training programs. Staff to collect and process samples is also needed to address the gaps revealed by the OTF scoping.

The need to recruit bioinformatics expertise into the federal workforce, provide more laboratory staff, and acquire more computing power are priorities identified by all surveyed Line Offices (Appendix II). The priority gaps in staffing and infrastructure noted by the OTF are shared across the federal agencies engaged with biological data (Stulberg et al. 2016, FTAC-MM 2015). One reason for this commonality is that the demand for 'omics information has increased computational demand, with biology rivaling the needs of astrophysics for computational capabilities (Stephens et al. 2015). This situation has been driven by greater sequencing capacity becoming available at increasingly lower costs. A decade ago, standard sequencing returned 1000 DNA base pairs; now millions are sequenced for the same cost. This revolution has provided an enormous opportunity, but it also poses a Big Data challenge. Terabytes of data can be generated by a single project, requiring that 'omics data be integrated into the agency programmatically.

New expertise and computational capacity are needed to ensure that emerging environmental data is expertly stored, managed, and analyzed. Otherwise, large quantities of valuable data could be generated, but not accessible as products and services used for management. Projections of computational capabilities necessary to support NOAA's 'omics activities are substantial (Appendix 1). The OTF survey revealed a need for internal infrastructure support because relying on extramural expertise and resources has been an inadequate strategy to address the gaps in bioinformatics expertise, data storage, and computing power.

## Investment in Bioinformatics and Analysis Capacity

To maintain scientific competitiveness and produce useful data products, NOAA must be current in its laboratory techniques, data pipelines and analysis, and computing hardware. Support should be provided for an 'omics infrastructure to collect, manage, process, archive, share, and analyze the large volumes of data. Capacity should be supplied centrally to NOAA's Information Enterprise and directly to NOAA's research laboratories and Science Centers. This capability is essential for successful implementation of long-term 'omics observations. Recommended focus areas include:

- Provide adequate data storage, processing, and analytical capacity to meet NOAA's requirements for throughput and efficiency.
- Access to modern, scalable cloud-based platforms.
- Develop permanent and publicly searchable archives for NOAA-generated genomic data and standardized metadata.
- Create central repositories of vetted analytical software to streamline and standardize workflows across labs.
- Leverage computational approaches such as machine learning and artificial intelligence to help interpret genetic variation and recognize relationships with environmental data.
- Procure infrastructure (e.g., sequencers) to collect 'omics data using a combination of in-house and commodity resources to support science applications (see Table 1) and to address fundamental issues such as improvement of databases.

The challenges found by the OTF are consistent with widely recognized priority challenges for research and monitoring in 'omics. Recognized challenges include the need for standardization of methods from sample collection through sample processing, more complete sequence databases that include metadata (contextual data that describe sampling and analysis parameters), and the need to improve quantification by both qPCR and next-generation sequencing (Goodwin et al. 2017). Incomplete databases affect all aspects of 'omics investigations. For example, the majority of the ~1 million marine species (not including prokaryotes) do not have reference sequences (Appeltans et al. 2012; Watts and Miksis-Olds 2018). Well-curated reference sequences are needed to interpret sequences obtained from the environment. Database gaps are a particular hindrance to long-term 'omics observations using eDNA approaches because typical bioinformatics approaches use binning algorithms to manage complexity. The final annotations (list of species or functional genes) are dependent on the state of the database at the time of analysis, which implies that time series data would require global re-computation as the databases grow. Alternative bioinformatics approaches (e.g., Deblur, DADA2) are being developed (e.g., Thompson et al. 2017), but more work needs to be done.

With regard to furthering eDNA applications, challenges with eDNA will need to focus on determining shedding and decay rates, the relation to species size, life cycle, and abundance, and on its fate and transport (i.e., its dependence on oceanographic and environmental variables) (see Appendix 1, NOAA Extramural Engagement in 'Omics; Appendix 2, and Hansen et al. 2017). The readiness of these ideas was clearly advanced at the 1<sup>st</sup> U.S. National Conference on Marine Environmental DNA, convened and hosted by The Rockefeller University, November 29-30, 2018

(Ausubel et al. 2018). Most notable in the Conference Report<sup>5</sup> is the Executive Summary, which simply reads: “eDNA works. Get going.”<sup>6</sup>

## Targeted and Coordinated Research

To take advantage of the full potential of ‘omics approaches, NOAA needs to invest in targeted research and standardizing molecular techniques to advance the state of the science. NOAA should leverage on-going observational programs that can support and benefit from routine collection of ‘omics samples. As outlined in the Introduction and in Table 1, priority areas for targeted investment include eDNA process studies, development of ‘omics time series, population and restoration genetics, aquaculture, and autonomous sampling and observation. Recommended research focus areas to harness efficiencies and transform research into solutions include:

- Include eDNA sampling and collecting in ocean exploration expeditions.
- Improved accuracy of resource analysis, fishery management recommendations, and assessment of ecosystem state through approaches that include: eDNA, close-kin mark-and-recapture, whole-genome analysis of commercially important species, and genetic identification and tracking of species that are difficult or impossible to identify visually (e.g., eggs, larvae, closely related species, microbes, pathogens, parasites).
- Use eDNA, or metagenomic, techniques to unravel trophic relationships through the study of food web structure and predator/prey relationships of fishery-important and protected species.
- Assessment of aquaculture species to ensure animal health and superior benefits for human consumption. Evaluate fitness, hatchery effects, and efficacy of outplanting efforts for endangered species such as corals, abalone, and salmon. Approaches include genomics (the study of organisms’ genomes), transcriptomics (the study of messenger RNA code and function), metabolomics (the study of cell metabolite chemical processes), and proteomics (the study of proteins and their biological functions).
- Improved reports and forecasts related to seafood safety, public health, and economic protection by utilizing molecular methods (eDNA qPCR or sequencing, immunoassays) and autonomous platforms to identify and track harmful algal blooms, pathogens, and invasive species.
- Forensics methods, such as genetic identification, to ensure that seafood products are accurately labeled and legally sourced (i.e., what they are and where they come from).
- Genetic approaches and eDNA to track and characterize populations of managed species that are difficult and expensive to assess by traditional methods, such as highly migratory species (HMS), and those that live in hard-to-reach habitats.

---

<sup>5</sup> [https://phe.rockefeller.edu/eDNAMarine2018/docs/MURU\\_eDNA\\_Conference\\_final\\_report.pdf](https://phe.rockefeller.edu/eDNAMarine2018/docs/MURU_eDNA_Conference_final_report.pdf)

<sup>6</sup> The eDNA component of ‘omics is greatly enhanced through the understanding of the subject and its ripeness as advanced by this workshop and its results. The first National Conference is a definite guide to NOAA and other federal agencies as to how to move forward through agency alignments and the National Ocean Partnership Program. These and other possible directions will be addressed in the Implementation Plan.

- Metagenomics to indicate water quality, ecosystem health, and bioprospecting potential.
- A suite of molecular methods to characterize the response of important species to ocean stressors (e.g., pollutants, disease, acidification, hypoxia, sedimentation/run-off, noise, warming) to design adaptive management practices.

## Partnerships

The fast pace of ‘omics advancement necessitates communication and coordination across NOAA’s lines and programs, across federal agencies, and engagement of the extramural research community. Recommendations similar to those found here have been proposed by other groups, such as the Microbiome Interagency Working Group<sup>7</sup> (MIWG). Such similarities are not surprising because microbiome research also relies on data derived from community sequencing and other genome-enabled technologies. *These commonalities support the idea that ‘omics science can be advanced by cross-line coordination and cross-agency interaction.*

To foster investment and to promote research transitions, ideas need to be shared across line offices and programs. A sample of already-existing groups working to enhance coordination is described in Appendix 1. Some of these represent grass-roots efforts to build collaborations while others are more directed. Examples of directed collaborations include the NMFS Strategic Initiative to apply eDNA and population genomics tools to support fisheries stock assessments, and an OAR’s ‘Omics grant program to the Cooperative Institutes (CIs) that supported protected resources research (corals) and sustainable fisheries needs. Ideas to improve internal communication include working groups to tackle specific problems, bioinformatics training workshops, meetings for ‘omics investigators, and sharing information via listservs. Other approaches to enhance NOAA’s ‘omics capabilities include collaboration with academic and private industry, including CI and Cooperative Science Center partnerships, Sea Grant, U.S. Integrated Ocean Observing System (IOOS), and research governed by Cooperative Research and Development Agreements (CRADAs).

Identified vehicles to promote opportunities and coordination in ‘omics across agencies include the Ocean Policy Committee’s Subcommittee on Ocean Science and Technology (SOST) and the National Oceanographic Partnership Program (NOPP). Establishment of an interagency ‘omics working group should be considered and include representation from NOAA, Bureau of Ocean Energy Management, Department of Energy (Hunter-Cevera et al. 2006), Environmental Protection Agency, U.S. Fish and Wildlife Service, National Institutes of Health, National Science Foundation, Office of Naval Research, Smithsonian Institution, U.S. Department of Agriculture, U.S. Army Corps of Engineers, and the U.S. Geological Survey.

Given the internal and external opportunities and collaborations, consideration should be given to establishing formal cross-NOAA line office and interagency advisory boards to foster partnerships, establish priorities, and joint evaluation and support of proposals and research.

---

<sup>7</sup> [https://www.nsf.gov/news/news\\_summ.jsp?cntn\\_id=245289](https://www.nsf.gov/news/news_summ.jsp?cntn_id=245289)

## Accelerate ‘Omics to Applications

Ensuring that advances in science and technology are transitioned to routine applications or operational use is a NOAA priority (Matlock et al. 2013; NAO 216-105B). The process of moving research to applications is often depicted as a pipeline or funnel, but the process is actually iterative. Critical elements for success include understanding mission requirements, assembling multi-disciplinary teams, and incorporating input from people slated to use the products. Harmful algal bloom (HAB) applications provide a good example of the research transition process, where decades of genetics research by NOAA and partners provided molecular diagnostics for bloom detection. Those assays were later engineered into automated platforms. Now, NOAA uses a commercially available instrument (e.g., MBARI’s Environmental Sample Processor<sup>8</sup> or “ESP”) to add necessary data for a HAB forecast in the northeast U.S., Gulf of Mexico, Pacific Northwest, Great Lakes, and other areas. IOOS Regional Associations have access to at least 14 ESPs operated by a range of partners for various applications. State and Tribal managers and other stakeholders use this information to help make decisions such as when to close or re-open commercial, recreational, and subsistence shellfish harvesting. In addition, MBARI and other scientists are exploring the use of the ESP both in its current form and future iterations as an eDNA collection platform.



There is a need to define the time to operational readiness for ‘omics applications. Table 1 identifies applications, their current status (concept, research, development, demonstration, applications, operations) and transition timelines. Timelines can be difficult to estimate because they are sensitive to technological innovation and resource investment, as well as end-user commitment to sustain operations. In general, Table 1 provides a timeline view of projects reviewed in the portfolio analysis (Appendix 1) and outlined in the “Investments and Development Areas” section. *To meet the transition timelines summarized here, there is a need to support workforce enhancement, increased analysis capacity and sequencing infrastructure, and targeted and coordinated research support.*

Of the use cases provided in Table 1, those rated with the highest readiness levels (~0-5 years to transition to routine use) include:

- Genetic species identification (versus traditional morphological analysis)
- Genomic-based population assessment using analysis of tissue
- eDNA abundance measurements of a particular species by qPCR
- eDNA biodiversity analysis by community sequencing

<sup>8</sup> <https://www.mbari.org/technology/emerging-current-tools/instruments/environmental-sample-processor-esp/>

- Organismal response and adaptation to variations in aquaculture rearing, diet, heat, acidification, etc. by analysis of RNA

Applications with intermediate readiness levels (~5-7 years to transition) include:

- Relative abundance
- Survivorship (e.g., salmon, turtles)
- Organism condition (e.g., aquaculture; sub-lethal stress)
- Ecosystem health assessments
- Environmental impact/damage assessments
- Stressor detection

Applications and methodologies available in the future (~10 or more years to transition) might include:

- Absolute abundance
- Restoration genetics
- Bioprospecting
- Engineering of algae to provide fish feed stocks
- Time series of ecosystem status

Laboratory and field research are needed to move the state of the science forward. In some cases, combining different technologies may be the best approach to overcoming technical hurdles and accelerating the transition of 'omics analyses to routine use. For example, acoustic analyses could be used to estimate fish biomass in conjunction with sequencing of eDNA for species composition and qPCR to quantify the amounts of DNA originating from target species. The concept of co-developing these complementary monitoring tools is reviewed by Watts and Miksis-Olds (2018).

Specific use cases define requirements and 'omics approaches. As such, Table 1 provides a detailed look at the applications listed in the "Investments and Development Areas" (Section 2 above). For example, measuring the differential expression of genes under varying environmental conditions allows insight into the health, reproductive condition, and growth rate of managed species. In turn, this information can be used to promote the sustainable management of resources. With regard to sustaining biodiversity, the discovery of corals more



## Research Highlight: Environmental DNA (eDNA)

Environmental DNA (eDNA) is one of many ‘omics approaches that promises to improve NOAA’s science. eDNA allows for the characterization of multiple species and genes from a single sample of seawater or sediment. In addition to investigating the base of the food web (bacteria and phytoplankton), eDNA can be used to detect higher organisms, such as fish or marine mammals, using samples of seawater (Hansen et al. 2018). Such “tissue-free” detection can be achieved by quantitative polymerase chain reaction (qPCR) amplification or by high-throughput sequencing of DNA that originated from sloughed or excreted animal cells (or whole plankton samples).

eDNA is discussed further in Appendix I. Briefly, the benefits of eDNA include:

- Non-destructive/invasive sampling (no nets, trawls, biopsies)
- Improved access (polar, deep, fragile, and other hard-to-reach habitats)
- Faster (better match with the time frame required to produce physical data)
- More comprehensive information (seawater can provide microbe-to-mammal surveys)
- More accurate identification of organisms at all life stages (from DNA sequencing)
- Sample archival in a cost-effective manner (e.g., frozen filters) and DNA sequences in public databases (e.g., NIH GenBank, Barcode of Life, etc.).

The NOAA ‘Omics Strategy includes a focus on furthering eDNA approaches. Promoting the transition of eDNA research into applications touches on the three main themes above: blue economy, efficient operations, and science & stewardship. This research area is a priority shared across multiple agencies and extramural partners, as reviewed in Appendix 1 and in the Introduction. A variety of applications for eDNA have been identified with a focus on sustaining fisheries, conserving protected species, protecting against harmful and invasive organisms, reducing labor and ship costs, and reducing the time between data collection and production of data products. Recommended steps to further eDNA applications include:

- Establish a baseline of biological inventory of the marine ecosystem in mapping and exploration of the ocean.
- Track invasive species, harmful algal blooms, aquaculture pathogens and parasites, migratory species, larval dispersal, and endangered populations.
- Investigate trophic interactions and improve understanding of food web dynamics, including connections between microbes and algal blooms or with animal stocks. Bottom-up drivers and prey selection by predators are the types of critical information needed.
- Provide remote sampling and eDNA archival of difficult or critical samples for species monitoring and population information, such as turtles and a variety of marine mammals.
- Deliver data on biodiversity to indicate ecosystem status or the efficacy of conservation measures.
- Provide economically viable sampling technologies to expand survey capabilities for data-poor fishery stocks.
- Augment “drone” platforms (UAV, AUV, gliders) for higher resolution biological detection, and automate environmental monitoring of aquaculture projects.

resistant to high temperatures and disease could jumpstart the effort to restore compromised coral reefs and consequently impact the thousands of jobs and billions of dollars associated with healthy reef ecosystems<sup>9</sup>. In the area of new discoveries and emerging markets, trait selection may produce improved aquaculture products, biosecurity could ensure that farming practices are safe and sustainable, and marine bioprospecting can yield new energy sources, biotechnology, and medicines. With regard to improving operations, it can take months or years to process biological data by traditional methods; ‘omics approaches offer faster and cheaper sample processing, which translates to faster public access to data and science-based conclusions.

## ‘Omics Broader Impacts

Economic activity directly associated with the ocean has been estimated to contribute \$304 billion to gross domestic product (GDP)<sup>10</sup>, with the U.S. fishing sector accounting for \$100 billion of GDP<sup>11</sup>. Coastal communities have been estimated to contribute 55.8 million jobs and produce more than \$8.3 trillion in goods and services<sup>12</sup>. In Florida alone, the estimated value of coral reef tract is \$8.5 billion and 70,400 jobs per year<sup>13</sup>. However, marine environments are facing an unprecedented set of global stressors (rising temperatures, acidity, nutrients, pollutants, etc.) that threaten ecosystems services. For example, microbial pollution at beaches can cost coastal communities up to \$37,000 per day (Scharff 2012), and seafood related infections from the *Vibrio* group of marine bacteria incur more than \$400 million per year in health costs (Rabinovici et al. 2004). The U.S. imports ~90% of seafood, with a \$14 billion seafood deficit in 2016<sup>14</sup>. Fundamental legal mandates for NOAA to address these issues include the Sustainable Fisheries Act of 1996, the Endangered Species Act, the Coral Reef Conservation Act, the Harmful Algal Bloom and Hypoxia Research and Control Act, the Marine Mammal Protection Act, and legislative advice to maintain healthy marine ecosystems and vibrant coastal communities.

‘Omics can address many of these concerns by providing an understanding of the genes and gene products that help marine organisms and ecosystems remain healthy. For example, safe and environmentally friendly aquaculture can foster job creation and help reduce the seafood import deficit. To this point, Lokman and Symonds (2014) state *“These ‘omics technologies...will contribute substantially to the future development of the aquaculture sector...It thus appears that ‘omics technologies may be a key tool for the responsible growth of intensive aquaculture in the 21st century.”*

---

<sup>9</sup> <http://dels.nas.edu/Report/Research-Review-Interventions/25279?bname=osb>

<sup>10</sup> National Oceanic and Atmospheric Administration (NOAA), Office for Coastal Management. 2019. "NOAA Report on the U.S. Ocean and Great Lakes Economy." Charleston, SC: NOAA Office for Coastal Management. Available at <http://coast.noaa.gov/digitalcoast/training/econreport.html>.

<sup>11</sup> <https://www.fisheries.noaa.gov/content/fisheries-economics-united-states-2016>

<sup>12</sup> <https://coast.noaa.gov/states/fast-facts/economics-and-demographics.html> accessed July 15, 2019

<sup>13</sup> Southwick Associates, 2003; Rabinovici et al. 2004; NMFS OST 2017; <https://floridakeys.noaa.gov/corals/economy.html>

<sup>14</sup> <https://www.fisheries.noaa.gov/national/aquaculture/us-aquaculture>

A number of nations are looking to implement ‘omics approaches such as eDNA, genetic species identification, microbiome characterization, and metabarcoding for applications that include monitoring of invasive species, determination of ecosystem status, and improved fisheries management (Goodwin et al. 2017; Hansen et al. 2018). In the European Union, implementation of such approaches is in concordance with the Marine Strategy Framework Directive (Hering et al. 2018). Bourlat et al. (2013) state *“The use of such approaches in next generation marine monitoring programs will help achieve the goals of marine legislation implemented world-wide...genomic technologies...can assist in producing rapid and rigorous information about ecosystem functioning, at a lower cost than traditional approaches.”* See Appendix 1 for examples of national and international collaborations working to advance the transition of ‘omics research to operations.

‘Omics approaches provide valuable opportunities across the agency to support the Blue Economy, promote efficient operations, and advance ocean science and stewardship. Recommended applications, outcomes, and overarching benefits include:

### **Support the Blue Economy**

- Advance aquaculture using ‘omics methods to improve animal health and management, animal products, and ecosystem monitoring.
- Protect health and associated local economies through improved harmful algal bloom and marine pathogen forecasts, including more accurate closures with regard to timing and geographic extent.
- Combat consumer fraud through genetic assessment (forensics) of seafood products.
- Help communities track the spread of invasive species and the efficacy of mitigation efforts by using eDNA to improve spatial and temporal monitoring at a lower cost.
- Preserve the ecological and socio-economic benefits of coral reefs by fostering more resilient coral species, and similarly kelp forest and other key communities.
- Sustain fisheries resources using ‘omics to inform stock assessments, including information about mixed stocks, food web/trophic dynamics, and life history linkages.
- Discover products for emerging markets (bioprospecting) and provide international leadership in the use of marine genetic resources.

### **Promote Operational Improvement**

- Provide faster access to comprehensive biological data compared to traditional sample processing methods (e.g., manual sorting and visual identification), eliminating backlogs and bringing the time scales of biological, chemical, and physical observations in sync.
- Find or develop autonomous samplers and vehicles that are cost-effective and increase sampling coverage, which are needed for NOAA model and forecast missions.
- Develop non-invasive and non-destructive biological sampling methods.

## Advance Ocean Science and Stewardship

- Understand and define the true biological distribution of organisms in the ocean and establish a biological baseline of ocean health.
- Protect highly migratory and vulnerable species through accurate definition of population structure and improved tracking.
- Accurately allocate incidental mortalities of marine species to source populations for responsive management actions.
- Use long-term ‘omics observations to better understand and characterize the influence of large-scale change on marine species and communities.
- Gain new insights into the life history and relatedness of marine species.
- Ensure that seafood is legally and sustainably harvested.
- Improve outcomes for hatchery-based recovery programs.

## *Summary and Next Steps*

‘Omics is a simple term to describe advanced tools in molecular biology. As is true in medicine, these tools applied to ocean science can improve NOAA’s products and services while controlling costs. High-throughput sequencing and bioinformatics analysis can be used to study the genetic mechanisms that help marine communities remain healthy. ‘Omics can be used to combat threats to marine ecosystem services, which contribute \$313 billion to GDP and generate 3.3 million jobs<sup>15</sup>.

NOAA has adapted ‘omics to address complex questions across multiple missions. Harnessing ‘omics technologies that improve our knowledge to better meet NOAA’s mission to manage, protect, and restore living marine resources is a long-term pursuit. Existing applications include combating toxic algae, harmful bacteria, and invasive species; increasing fisheries and aquaculture productivity while managing impacts to vulnerable species and habitats; and helping corals, which provide essential fish habitat and fuel tourist economies. Genetic measurements can be a sentinel for ecological status by measuring biodiversity, food web interactions, and organism stress. Overall, ‘omics approaches are valued for their promise to improve evaluation of baseline and changed ecosystem function, to provide new understanding of biodiversity that will aid exploration and stewardship, to help protect health and economies by allowing stakeholders to manage and mitigate contamination issues, and to engage in international, academic, and commercial partnerships.

There are, however, immediate areas for development to enhance current efforts and maximize the utility of ‘omics tools. This white paper describes priority areas requiring NOAA’s investment and development, in particular:

- Enhancement of a well-trained workforce
- Investment in bioinformatics infrastructure and analysis capacity
- Prioritization of targeted and coordinated research

---

<sup>15</sup> <https://coast.noaa.gov/digitalcoast/data/>, see Economics: National Ocean Watch - GDP

NOAA recognizes effective management and protection of ocean resources requires leveraging modernized technology including investment, utilization, and implementation of ‘omics technology. With proper training and resources, NOAA can accelerate understanding of how marine ecosystems adapt to environmental change at local, regional, and global scales. The aim is to better meet NOAA’s mission to manage, protect and restore living marine resources in support of NOAA’s blue economy priority.

A systematic approach is recommended to support a spectrum of activities from capacity building in bioinformatics (training, research, recruitment, and student internship opportunities), integration of ‘omics into assessments tethered to NOAA mandates, research-to-application projects to reduce operational costs, and applied genomic research linked to NOAA products and services.

The concepts outlined here have been used to inform the NOAA ‘Omics Strategy and the NOAA ‘Omics Implementation Plan. Briefly, the NOAA ‘Omics Strategy lays out the following Goals and Objectives:

### NOAA ‘Omics Strategy Goals and Objectives

Goals	Objectives
<p><b>Goal 1:</b> <b>Enhance infrastructure to meet the analytical demands of ‘omics data.</b></p>	<p><b>LABS, SHIPS, AND VEHICLES</b> <i>Objective 1.1.</i> Provide adequate laboratory space in facilities and ships to collect, process, and store samples for ‘omics analyses, and increasingly leverage UxS (unmanned systems) for data collection (see <a href="#">NOAA Unmanned Systems Strategy</a>).</p>
	<p><b>COMPUTING POWER and STORAGE</b> <i>Objective 1.2.</i> Procure the analytical and computational infrastructure needed to generate, analyze, and manage massive ‘omics data sets, and increasingly leverage the commercial cloud for computation and data storage (see <a href="#">NOAA Cloud Strategy</a>).</p>
	<p><b>BIOINFORMATICS and DATABASES for ‘OMICS TIME SERIES</b> <i>Objective 1.3.</i> Expand the databases that identify genetic sequences and develop bioinformatics tools needed to manage and interpret time series data, including impacts of large-scale environmental change through biodiversity monitoring.</p>
	<p><b>SHARED REPOSITORIES</b> <i>Objective 1.4.</i> Create a central repository to share protocols, standards, and house bioinformatics pipelines to support a community of practice across laboratories and programs.</p>
	<p><b>MACHINE LEARNING AND AI</b> <i>Objective 1.5.</i> Leverage computational approaches such as machine learning and artificial intelligence (see <a href="#">NOAA Artificial Intelligence Strategy</a>) to help interpret genetic variation and recognize relationships with environmental data.</p>

<p><b>Goal 2: Execute 'omics research tar- geted to support and advance the Blue Economy</b></p>	<p><b>BIO-SURVEILLANCE</b> <i>Objective 2.1.</i> Improve detecting and monitoring of harmful algal blooms, toxins, pathogens, and invasive species to protect health and coastal economies.</p>
	<p><b>SEAFOOD FORENSICS</b> <i>Objective 2.2.</i> Support consumer protection and sustainable fishing practices by using genetic analysis to identify fraudulent and illegally sourced seafood products.</p>
	<p><b>SUSTAINABLE AQUACULTURE</b> <i>Objective 2.3.</i> Foster the development of aquaculture by using 'omics to optimize animal health, yield, and product characteristics while supporting safe and sustainable farming practices.</p>
	<p><b>FOOD WEBS, FISHERIES, and PROTECTED RESOURCES</b> <i>Objective 2.4.</i> Sustain fisheries resources and protect vulnerable species using 'omics to increase the breadth, depth, and throughput of information used to evaluate target populations' structure and distribution, generate indices of abundance, and characterize the food webs that support them.</p>
	<p><b>BIODIVERSITY AND BIOPROSPECTING</b> <i>Objective 2.5.</i> Advance the exploration of biodiversity and bioprospecting to discover natural products that may have medical or other commercial value and provide international leadership in the use of marine genetic resources while protecting biodiversity.</p>
<p><b>Goal 3: Accelerate transition of 'omics research into operations</b></p>	<p><b>UNDERSTAND AND FULFILL MISSION REQUIREMENTS</b> <i>Objective 3.1</i> Conduct field trials to define operational requirements, calibrate 'omics approaches with traditional methodologies, and clarify design specifications to accelerate production of validated approaches.</p>
	<p><b>STANDARDIZED, INTEROPERABLE AND AVAILABLE DATA</b> <i>Objective 3.2.</i> Promote a unified approach to sample and metadata collection, sample processing, and data deposition in publicly searchable archives to promote interoperability and time series establishment</p>
	<p><b>INDICATORS AND ACTIONABLE EVIDENCE</b> <i>Objective 3.3.</i> Develop and integrate 'omics ecosystem indicators into reports, models, and forecasts to benefit seafood safety, public health, and economic protection.</p>
	<p><b>COMBINE TECHNOLOGIES TO ACCELERATE SUCCESS</b> <i>Objective 3.4.</i> Combine 'omics with existing and emerging technologies to synergize the strengths of individual approaches and thus hasten the innovation of operations.</p>
	<p><b>UTILIZE R2X PROCESS</b> <i>Objective 3.5.</i> Develop transition plans with NOAA Line Office Transition Managers (LOTMs) to outline steps for technology transfer and provide incentives and support for 'omics R2X.</p>

<p><b>Goal 4:</b> <b>Expand partnerships to advance 'omics research and applications across the agency</b></p>	<p><b>INTERNAL COMMUNICATION AND CHAMPIONSHIP</b> <i>Objective 4.1.</i> Establish a NOAA Executive Committee, chaired by the Chief Scientist, to guide the 'Omics Working Group (OWG) to share information opportunities, and promote the priorities outlined in this strategy across the agency.</p>
	<p><b>ENGAGE USER COMMUNITIES</b> <i>Objective 4.2.</i> Engage existing national and international groups working to enhance 'omics technology improvement, standardization, long-term observations, and data and sample archival.</p>
	<p><b>INTERAGENCY FUNDING OPPORTUNITIES</b> <i>Objective 4.3.</i> Prioritize 'omics research in existing interagency funding opportunities to advance 'omics research and development.</p>
	<p><b>NATIONAL AND INTERNATIONAL ENGAGEMENT</b> <i>Objective 4.4.</i> Foster coordinated and collaborative projects across agencies and internationally to advance 'omics applications.</p>
	<p><b>TECHNOLOGY TRANSFER PARTNERSHIPS</b> <i>Objective 4.5.</i> Build and sustain partnerships with the private and academic sectors using existing vehicles to encourage engagement with federal 'omics research and development and to increase the potential for commercialization.</p>
<p><b>Goal 5</b> <b>Promote workforce proficiency in 'omics</b></p>	<p><b>NEEDS ASSESSMENT</b> <i>Objective 5.1.</i> Conduct a baseline needs assessment to inform goal implementation.</p>
	<p><b>TRAINING</b> <i>Objective 5.2.</i> Provide training for 'omics data collection and bioinformatics analysis to increase expertise within the current workforce.</p>
	<p><b>RECRUIT AND RETAIN</b> <i>Objective 5.3.</i> Recruit and retain information technology (IT) professionals and scientists with bioinformatics expertise to address current gaps in the ability to analyze and provide biological or environmental context to sequence data.</p>
	<p><b>DEVELOPMENT OPPORTUNITIES</b> <i>Objective 5.4.</i> Develop opportunities for job details in laboratory facilities to provide career development for staff, interns, and fellows, and to promote 'omics projects and data integration.</p>
	<p><b>SHARE EXPERTISE</b> <i>Objective 5.5.</i> Focus assignments in the NOAA Rotational Assignment Program (NRAP) to target offices where a cross-pollination of 'omics expertise would raise overall proficiency.</p>

**Table 1. Implementation Status of ‘Omics Technology Applications**

<b>Approach</b>	<b>Application</b>	<b>Status</b>	<b>Timeline to Operations (years)</b>
Environmental DNA (eDNA) - water sampling	Species monitoring (e.g., toxic algae, invasive species, bacterial pathogens, or bacterial load to prepare communities, inform forecasts, and devise mitigation strategies)	Demonstration	presence/absence: 0-5 (community sequencing) quantification: 0-5 (qPCR, single target)
	Determine and monitor prey usage of protected marine mammal species	Demonstration/Operational	0-5 depending on species; currently used for Southern Resident Killer Whales
	Biodiversity assessments to determine ecosystem health and response to global change and to achieve biodiversity goals	Research/Demonstration	individual assessments: 2-7 maintenance of time series: 5-10
	Assess currently unmonitored fish/shellfish stocks due to crypticism and/or sampling restrictions	Research/Demonstration	3-5 or 5-10, depending on stock and geographic location
	Inform stock assessments through additional temporal-spatial data; close-kin mark-and-recapture for population estimates	Research/Demonstration	3-10+, depending on stock and geographic location
	Monitor for water quality along coastlines by measuring bacterial load	Research	0-3, depending on application and target
	Surveillance of aquaculture pathogens and monitor the environmental impact of aquaculture	Research	3-10, depending on resource allocation
	Monitor protected species to conserve populations	Research	1-10+, depending on level of population resolution required
Organism DNA – tissue sampling	Identify endangered species and populations using genetic markers to establish effective management plans for conservation	Operational	Currently Used
	Identify illegally mislabeled seafood	Demonstration/Operational	Operational for some species but still small in scale
	Select for traits that make organisms resistant to changing environmental conditions (i.e. thermal tolerance, disease resistance)	Research/Demonstration	Demonstrated for some coral species
	Select for traits desirable for aquaculture (i.e. time to sexual maturity, size, quality of meat)	Research	2-10, depending on stock
	Understand and monitor species sensitivity to ocean change	Research	10+
Proteomics	Monitor and improve organism health for aquaculture and recovery of species	Research	10+
	Understand and monitor species sensitivity to ocean change	Research	10+
Metabolomics	Bioprospecting for new medical drugs or other commercially valuable compounds	Research	10+
	Understand and monitor species sensitivity to ocean change	Research	10+



# APPENDIX I.

## PORTFOLIO ANALYSIS

The following provides a summary of information collected by the OTF regarding the current portfolio of 'omics related activities, capacities, and collaborations across NOAA.

### National Marine Fisheries Service (NMFS) Activities in 'Omics

A number of NMFS laboratories utilize 'omics techniques to carry out NOAA missions. Genetic and genomic investigations, for example, have been carried out for decades. In addition, most NMFS laboratories are exploring newer approaches, such as eDNA analysis. The analysis of environmental DNA (eDNA) is proving to be a powerful tool for species detection in diverse environments (Bohmann et al. 2014, Rees et al. 2014, Thomsen et al. 2015) and shows great potential in cost effective mapping and monitoring of biodiversity. In general, eDNA analysis detects species presence in aquatic environments with equal or higher sensitivity than conventional sampling methods (Bergman et al. 2016, Erickson et al. 2016, Port et al. 2016, Schmelzle et al. 2016 Thomsen et al. 2016). In a nearshore marine environment, the eDNA approach generally outperformed nine conventional methods (bottom trawling, beach seine, multi-mesh gill nets, angling, push nets, fish pots, fyke nets, day-snorkeling and night-snorkeling), detecting 15 commonly occurring fish species using just three samples of 0.5 L of seawater. Most of the conventional methods detected fewer species than eDNA analysis and none found species which were not detected by eDNA (Thomsen et al. 2012). Recent studies have also found positive correlation between eDNA concentration in water and species biomass (Lacoursière-Roussel et al. 2016, Yamamoto et al. 2016), suggesting that quantification of eDNA may, under certain conditions, be used for fisheries management. In addition, eDNA studies are non-invasive and less reliant on rapidly dwindling taxonomic expertise, under-scoring the large potential of the approach for solving many pressing issues for marine science and management (Goodwin et al. 2017).

## NOAA FISHERIES

### NOAA Fisheries Activities that can utilize 'Omics

- Species abundance, ID, distribution (manage ecosystem resources, stock assessments)
- Stock structure of protected species (populations to conserve)
- Early warning of harmful algae (ecological forecasting)
- Aquaculture applications (environmental impacts, food safety & security)
- Forensics (IUU fishing)



Research activities by NMFS and Cooperative Institute partners that heavily use 'omics technologies include: analysis of species relative abundance in trawls (e.g., surveys and stock assessments), species abundance in un-trawlable and near-shore habitats, distribution of non-native species, species identification of deep sea corals, taxonomy and stock structure of managed and protected vertebrate species, early warning systems/ecological forecasting (for example harmful/toxic algal blooms for seafood safety applications), studies of diet composition, and a variety of aquaculture uses. A brief overview of projects is provided below, and more information can be found in Werner et al. (2018):

- Population genomics for a variety of species to delineate stock structure and breeding populations, including those originally thought to be panmictic, to improve stock assessments and management, including conservation and recovery. Applications include commercially valuable fish, protected turtle and marine mammal species, and identification of animals in anthropogenic and natural mortality events.
- Next-generation DNA sequencing approaches to inform the stock structure of salmon bycatch in federally managed fisheries on the US west coast (from California to Alaska), and for sharks caught as bycatch in Alaska long-line fisheries.
- Use of the 2nd Generation Environmental Sample Processor (ESP), to provide near real-time data delivery of ecosystem status in Puget Sound and in the Pacific Ocean off the coast of Washington, with regard to harmful algae and algal toxins. This work represents a successful transition of research to data products.
- Development of eDNA survey methods to quantify near-shore fish communities with focus on salmon, herring, and smelt species. Methods include quantitative polymerase chain reaction (qPCR) techniques, massively parallel DNA sequencing, and development of a new statistical framework to provide field estimates of the relationship between species abundance and eDNA. Survey of Pacific hake (Pacific whiting) to compare eDNA with trawl data (acoustic and fishing) to augment survey data for fisheries stock assessments is ongoing.
- Measurement of shedding and degradation rates of eDNA from adult black sea bass in relation to environmental conditions such as temperature and pH under controlled laboratory conditions. Monitoring eDNA to assess the effects of cultured caged oyster on finfish diversity in Long Island Sound. Developing an eDNA Toolkit for Northeast Diadromous Fishes.
- eDNA analysis to provide non-invasive, non-destructive, species-level identification of deep-sea corals (which are difficult to identify by sight); work includes forensics applications. Application of eDNA to monitor the distribution on non-native (invasive) walleye in Lake Washington (Washington State).
- Applications of genomics, epigenetics, transcriptomics, and metabolomics to guide sustainable marine aquaculture. Efforts include identification of desirable genetic traits (growth rate, disease resistance, etc.), development of improved diets, understanding mechanisms of fitness loss in hatchery-raised fish, and the effects of large-scale environmental change on fisheries resilience. For example, analysis of sablefish sex differentiation genes led to a method to create all-female progeny using RNA-silencing; female sablefish grow faster and reach market size quicker, a desirable trait for finfish aquaculture.
- Microbiome analysis of gut contents of sablefish fed different alternative and/or plant-based feeds, in order to better understand feed conversion, needed for sustainable aquaculture.
- Applications in genomics, transcriptomics, and metabolomics to evaluate the response of Dungeness crab, red king crab, geoduck clam, and salmon to ocean acidification.



- Next-generation DNA sequencing applications include identification of epiphytic microorganisms associated with sugar kelp to support northeast aquaculture; stomach contents for cod, seabirds, and marine mammals; ichthyoplankton and zooplankton species identification and diversity in support of ecological monitoring surveys.
- Use of eDNA to differentiate stock structure of harbor porpoise to overcome difficulties in obtaining biopsies in southeast Alaska. Use of genomic sequencing to establish baseline genomic variation and information for potential recovery of the world's most endangered marine mammal, the Vaquita porpoise.
- Genomic and phylogenomic comparisons of multiple strains of the human pathogen *Vibrio parahaemolyticus* led to a better understanding of specific sequence types or clades capable of causing disease, including first genomic analysis of what is now a new pandemic strain, thereby improving forecasts for the pathogen in shellfish aquaculture.

### **Bioinformatics Analysis and Needs in NOAA Fisheries Science Centers**

Enhancement of bioinformatics and IT enhancement to address bottlenecks in data analysis and interpretation include NOAA Fisheries computer cluster resources at Northwest Fisheries Science Center (NWFSC) and Southwest Fisheries Science Center (SWFSC). These specialized systems are local and can be quickly overcome by local science center needs if scalability is not accounted for in the original installation of these resources. For a more complete description, see "Current Bioinformatics Infrastructure" in Appendix 2. Use of NOAA High-Performance Computing Cluster (HPCC) resources for bioinformatics has been attempted, but were found to be lacking due to highly parallel but low RAM configuration of each node in the HPCC systems (genome and transcriptome assemblies require high memory nodes - see Appendix 2).

### **NMFS Strategic Initiative on 'Omics (FY19-FY24)**

In FY19, the NMFS Science Board approved a 5-year effort dedicated to advancing the use of 'Omics technologies in the support of the fisheries stock assessment enterprise. NMFS Science Centers involved in the initiative are using eDNA and population genomics methodologies to augment or advance data streams that feed important stock assessments on both east and west coasts (groundfish and coastal pelagic species). In addition, eDNA methods to eDNA/meta-genomic techniques to unravel trophic relationships through the study of food web structure and predator/prey relationships of fishery species are being undertaken. In addition, an investment in new shared computational resources and advanced bioinformatics training has been made and is showing great interest and promise across the agency.

## Oceanic and Atmospheric Research (OAR) Activities in ‘Omics

### Aligned OAR ‘Omics Laboratories

The Aligned OAR ‘Omics Laboratories group currently includes the Atlantic Oceanographic & Meteorological Laboratory (AOML), the Great Lakes Environmental Research Laboratory (GLERL), the Pacific Marine Environmental Laboratory (PMEL), and associated Cooperative Institutes (CIs). Supported activities in ‘omics focus on promoting the Blue Economy, technologies to improve NOAA’s operational efficiency, and transitioning research into information and products sought by management. Plans are coordinated among AOML, GLERL, and PMEL to best build OAR’s molecular and bioinformatics capacity. Expertise is distributed across the laboratories, promoting collaboration among the laboratories and CIs.

Science applications, tools, and approaches shared broadly across the laboratories include mechanistic studies encompassing food web dynamics and organism response to ecosystem stress, establishment of eDNA/microbiome time series to capture critical trends, and development of autonomous platforms to augment collection of ‘omics data. In addition to shared themes, specialization exists to address specific regional needs such as promoting resilience and restoration of shallow-water corals (AOML), combating harmful algae/invasive species in the Great Lakes (GLERL), and characterizing Pacific and Arctic food chain and deep-sea vent resources (PMEL). All efforts are underpinned by building the capacity and expertise needed to successfully implement the NOAA ‘Omics Strategy.

The vision statement for the Aligned OAR ‘Omics Laboratories is:

*Through aligned efforts, OAR’s Omics Laboratories will build a central backbone of resources, lead bioinformatics development, and spark the next generation of ecosystem assessments to address critical regional and management needs.*

The Aligned OAR ‘Omics Laboratories conduct research that is informed by the NOAA ‘Omics Strategy. Common themes to implement the NOAA ‘Omics Strategy include:

- **Build needed capacity** across the laboratories to collect, process, maintain, and analyze needed ‘omics data, including molecular and bioinformatic protocols, processes, expertise, and infrastructure.
- **Conduct mechanistic research** to address critical knowledge gaps and inform strategic actions related to reef restoration, causes and consequences of harmful algae impacts of invasive species, and fishery resource fluctuations.
- **Capture trends** that inform management by establishing and maintaining time series of Ecosystem and Community ‘Omics (ECO).
- **Transition tools and discoveries** into operational practices and products to help understand and predict ecosystem and resource response.

The table below maps selected FY20 research projects to the themes outlined above. These projects involve multiple partners, which are not listed here due to space considerations. These

projects map to priority applications described in the NOAA 'Omics Strategy including fisheries management, food and water safety, species and habitat conservation, aquaculture development, and natural products discovery.

### Selected Projects Mapped to Shared Themes

Lab Lead	Capacity	Mechanistic	Time Series	Transition	Brief Project Description
PMEL	X	X	X	X	Pacific coast community relationships to env. conditions
PMEL	X	X	X	X	Alaskan and Arctic planktonic characterization
PMEL	X	X		X	Chemosynthetic community connectivity
GLERL	X	X		X	Rise & demise of HABs
GLERL	X			X	Toxin detection for 3G ESP
GLERL	X		X		eDNA and metabarcoding: invasive mussel impact
GLERL	X		X		Great Lakes 'omics database
AOML	X	X		X	Molecular insights from corals in extreme envs.
AOML	X	X		X	Molecular mechanisms behind coral resilience
AOML	X		X	X	Cooperative advancement of eDNA
AOML	X		X	X	Developing autonomous 'omics
AOML	X		X	X	Bioinformatics to support the 'omics roadmap
AOML	X		X	X	Environmental DNA for fisheries monitoring
AOML	X				Supporting 'omics
AOML		X		X	Heat tolerant coral genotypes for restoration
AOML	X	X			Bioinformatics for microbiome studies
AOML	X	X			Pathogens with red tide

The vision and strategy of project implementation for the Aligned OAR 'Omics Laboratories group was created through the active engagement of 25 NOAA federal and CI employees. The group holds meetings and workshops to share research findings and technical recommendations and to facilitate coordinated implementation and advancement of 'omics methods. More detailed information about activities within individual laboratories is also provided below.

## Atlantic Oceanographic & Meteorological Laboratory (AOML) 'Omics Program

AOML strives to facilitate the protection of natural resources and sustainable economic growth through better monitoring and forecasting of ecological conditions to inform decision making at local, regional, and national levels. The laboratory, and particularly AOML's Ocean Chemistry and Ecosystem Division, employs a cross-disciplinary and collaborative approach in conducting its research programs. Core elements of the AOML mission include obtaining reliable measurements and stewardship of those data, development and deployment of new sampling methods and analysis tools, and implementation of long-term, consistent environmental observation programs. AOML's coastal ocean and ecosystem research has studied organisms from small microbes to large marine mammals and areas from the intertidal zone to deep-sea hydrothermal vents. Research is conducted through a suite of sustained monitoring projects, targeted process studies, and ecological models. Applications include helping corals, which provide essential fish habitat and fuel tourist economies; promoting fisheries productivity while managing impacts to vulnerable species and habitats; and combating toxic algae and harmful bacteria. The unifying goal is to improve our management of coastal ecosystems, thereby maximizing ecosystem health and economic yield.



The 'omics program at AOML works to promote coral resilience, develop and transfer emerging technologies, advance 'omics for fisheries and microbiome applications, and foster the bioinformatics and infrastructure capabilities upon which all 'omics research and operations rely. This work engages coral, fisheries, and microbiome experts across the agency and via international engagement. Portfolio activities are organized into the following themes:

***Coral Restoration 'Omics:*** Coral reefs are in crisis globally. Coral reefs provide billions of dollars to the U.S. economy annually through fisheries, tourism, and storm protection. In southeast Florida alone, coral reefs are estimated to generate \$4.4 billion in local sales, \$2 billion in local income, and 70,400 full and part-time jobs. Widespread coral mortality in recent years has led to a decline in coral cover and erosion of reef framework structures. In the Caribbean, coral cover has declined by roughly 80% in the last three decades, sending this marine resource in a downward trajectory with strong ramifications for ecosystem services and the survival of species.

'Omics can serve NOAA's mission to conserve important marine resources by increasing the efficacy and efficiency of reef restoration and management. AOML is working to identify resilient coral genotypes and to isolate the molecular mechanisms by which resilience is conferred. Field and laboratory research are used to identify and understand coral genotypes and microbiomes associated with resistant phenotypes (heat, acidification, and disease). Information regarding the strengths and weaknesses of a given genotype can inform management decisions regarding where and which individuals should be out-planted to maximize restoration success. Identifying the molecular underpinnings of this resilience/susceptibility will help to rank and select candidate genotypes for informed and effective restoration.

**Advancing Environmental DNA (eDNA):** Mandates for integrated, ecosystem-based monitoring drive a global need for low-cost bioindicators of marine ecological quality. The microbiome forms the base of the food web and controls cycles of carbon, nutrients, oxygen, metals, and toxins. As we have learned about the diversity of microorganisms and their associated biogeochemical processes, our view of ocean ecosystems has been expanded to appreciate the relevance of microbes to all aspects of life, including ocean resiliency and marine resource management. In addition to microbes, assessments of higher trophic levels, such as fish or marine mammals, can be gleaned from environmental DNA (eDNA) – DNA from filtered seawater via capture of sloughed or excreted cells. eDNA is collected from seawater or sediment instead of from tissues so that samples can be collected without animal capture, tissue processing, or trawling through sensitive habitats.

**Autonomous ‘Omics:** Molecular methods are being integrated into autonomous platforms. NOAA and partners are field testing a prototype long-range autonomous underwater vehicle (LAUV) that can search for oceanographic features and filter water for genetic analysis using a 3<sup>rd</sup> generation Environmental Sample Processor (3G ESP). This prototype instrument can be deployed from dock or small boat, allowing rapid response, and increased sample coverage while combating rising ship-time costs, which is important for NOAA modeling and forecast missions.

**Fisheries ‘Omics:** AOML is working to transition ‘omics research into fisheries management applications. An overarching hypothesis being tested by a suite of projects is that genomic information integrated into fisheries management will improve the decision-making process and thus the sustainability of fisheries.

**Building ‘Omics Workforce and Infrastructure:** Genome-based techniques improve our ability to characterize and monitor ecosystems. However, our ability to supply bioinformatics expertise has not kept pace with the generation of sequence data, creating a data and analytical backlog and hindering transition of data collected into actionable information. To address this gap, AOML is dedicated to modernizing our workforce and infrastructure, which is important to the success of all ‘omics projects.

More detailed information on activities within these themes is provided below.

*The Experimental Reef Laboratory:* The Experimental Reef Laboratory<sup>16</sup> physically located at the University of Miami’s Marine Technology Life Science Seater facility, was completed in September of 2016. This unique experimental facility allows unprecedented control of seawater temperature and chemistry, with control of pH to 0.05 units. Scientists measure how coral organisms respond at the molecular level (DNA and RNA) under present and possible future conditions. Studies work to quantify responses to heat, light, and acidification, identify mechanisms that confer resistance to stress, and evaluate methods for treating coral disease. Work aims to understand how environmental conditions contribute to coral susceptibility/resilience to allow managers to identify potential species-specific refugia and restocking strategies to provide ecosystem recovery.

---

<sup>16</sup> <https://www.aoml.noaa.gov/keynotes/PDF-Files/Jan-Feb2017.pdf>

*Restoration Genetics:* AOML is working to identify resilient coral genotypes and to isolate the molecular mechanisms by which resilience is conferred. In-depth field research of coral bleaching patterns has revealed that some corals are more resistant to bleaching from heat stress. They are associated with genetic signatures that are adapted/acclimatized to +1.0°C hotter than others of the same species. These results provide hope of using resistant genotypes to ensure survival of endangered species and restoration of reef habitat. AOML studies corals from marginal coral environments (e.g., Port of Miami) that have nursery genotypes identified to be resistant to disease through experimental assays, as well as wild genotypes identified to be resistant to heat stress and bleaching. The goal is to identify pathways and markers for resilience via transcriptomics and proteomics of wild and experimentally manipulated corals and to develop stress assays to rapidly evaluate and *a priori* characterize bleaching susceptibility.

*Genetics of Disease Resistance:* AOML is working to address the devastating consequences of disease spreading through South Florida and the Florida Keys at the molecular level. Scientists are modeling regional disease connectivity and disease flow pathways, improving the predictability of the spread of disease. The aim is to inform outplanting of disease-resistant genotypes, develop diagnostic tools to identify resistant individuals, and monitor the long-term viability of less-susceptible coral holobionts. Furthermore, microbial source tracking is being used to track land-based sources of pollution, and DNA sequencing is being used to investigate the biodiversity of coral microbiomes<sup>17</sup>.



*Emerging Technologies: Improving Operations through Automated Sampling, eDNA, and Metabarcoding:* ‘Omics can improve NOAA’s operational efficiency resulting in reduced operating costs and increased public access to data. Work focuses on molecular methods and technologies that reduce sample backlog, environmental impact, and costs compared to conventional sampling and sample processing methods.



*NOAA-MBARI Collaboration:* Molecular methods are being integrated into autonomous platforms. In collaboration with the Monterey Bay Aquarium Research Institute (MBARI), NOAA is field testing a prototype long-range autonomous underwater vehicle (LAUV) that can search for oceanographic features and filter water for genetic analysis using a 3<sup>rd</sup> generation Environmental Sample Processor (3G ESP). This prototype instrument can be deployed from dock or small boat, allowing rapid response, and increased sample coverage while combating rising ship-time costs, which is important for NOAA modeling and forecast missions. Major progress with the 3G ESP/LRAUV (“eAUV”) includes increased sampling capacity, and the

---

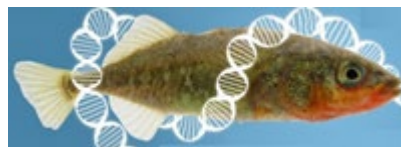
<sup>17</sup> <https://www.aoml.noaa.gov/keynotes/PDF-Files/Mar-Apr%202017.pdf>



ability to perform on-board toxin detection. Several missions have been conducted in the California Current marine ecosystem<sup>18</sup>. In addition, researchers with MBARI, NOAA (AOML and GLERL), and the Cooperative Institute for Great Lakes Research tested the eAUV in the freshwaters of Lake Erie<sup>19</sup>.

*Subsurface Automatic Samplers (SASs) for the Collection of eDNA:* AOML is working to develop inexpensive auto-samplers for the collection and preservation of eDNA to advance the application and utility of 'omics-based technologies in marine science. There is a critical need for small, relatively inexpensive, simple, and easy to handle technology that can automatically filter water and preserve the DNA on filters for extended periods of time at ambient conditions and temperatures. This work builds on open-source subsurface automatic samplers (SAS)<sup>20</sup> developed in AOML's Advanced Manufacturing and Design Lab (AMDL). The SASs will serve as a low-cost solution for sampling eDNA during periods when field collection is logistically or financially unfeasible.

*Environmental DNA (eDNA):* eDNA is a way to detect higher trophic levels, such as fish or marine animals, using only seawater or sediment samples. Instead of tissue samples, the detected DNA comes from cells that have been sloughed or excreted into the environment so that samples can be collected without animal capture, tissue processing, or trawling through sensitive habitats. NOAA is investigating whether eDNA can provide information about commercially important and protected species without the need to process tissues or trawl through sensitive habitats, with focus on fish and turtles.



*Metabarcoding:* Manual sorting and visual identification of larval fish and eggs is used routinely by NOAA. However, the process is labor intensive and slow, causing data backlog measured in years. In collaboration with the Southwest Fisheries Science Center (SWFSC), ichthyoplankton metabarcoding is being developed to enable bulk analysis of samples to increase efficiency, reduce costs, and decrease dependence on specialized morphologic expertise. Work is also being done on ethanol-preserved samples with the goal of using eDNA from the bulk ethanol to ascertain sample information without having to destroy archived tissues samples. Advancing this capability would be useful to a number of NOAA laboratories and programs (CalCOFI, SEAMAP, and others).

*Fisheries Applications:* AOML is working to transition 'omics research into fisheries management applications. Our ability to effectively manage fisheries is limited by our understanding of fishery populations and their dependence on environmental conditions. An overarching hypoth-

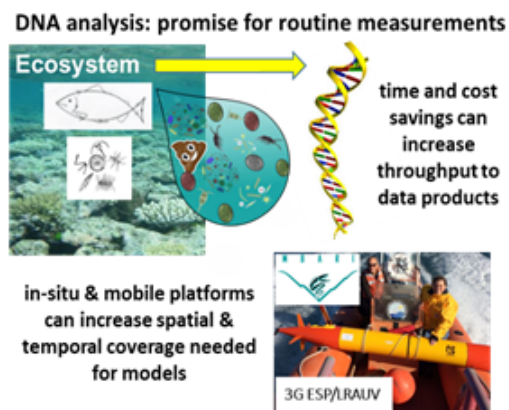
---

<sup>18</sup> <http://www.aoml.noaa.gov/keynotes/PDF-Files/May-June%202017.pdf>

<sup>19</sup> <https://research.noaa.gov/article/ArtMID/587/ArticleID/2382/NOAA-and-partner-test-unmanned-vehicle-to-detect-harmful-algal-blooms-in-Lake-Erie>

<sup>20</sup> <https://www.coral.noaa.gov/accrete/sas/>

esis being tested by a suite of projects is that including genomic information into fisheries management improves the decision-making process and thus the sustainability of fisheries. How genomic information affects decisions will be evaluated using the Management Strategy Evaluation framework adopted by the National Marine Fisheries Service. AOML and the Southeast Fisheries Science Center (SEFSC) have formed a collaborative with a clear path to transition results into fisheries management plans for the following projects: 1) using DNA Single Nucleotide Polymorphisms (DNA SNPs) to differentiate between eastern and western stocks of Atlantic bluefin tuna to inform stock assessments, and 2) using Restriction site Associated DNA Sequencing (ddRAD-Seq) to access the population structure of king mackerel to address pressing uncertainties affecting the stock assessment of this species. In addition, empirical models are being used to assess whether microbiome data (sometimes termed “eDNA”) improves the ability to predict abundances of fish and to understand trophic and habitat interactions in collaboration with 3) a Marine Biodiversity Observing Network (MBON) project, and 4) the NCOG project (see above), in collaboration with the Southwest Fisheries Science Center (SWFSC).



**Microbiome Applications:** Lower trophic levels are foundational to an ecosystem because they form the base of the food web and control the cycles of carbon, nutrients, oxygen, metals, and toxins. For decades, biological oceanographers have used traditional measurements, such as chlorophyll or total biomass, to describe this essential component of the earth system, but with limited predictive success for the status of ecosystem health or stocks of economically critical species. This gap may exist because traditional measurements are unrefined, with little to no detailed information, particularly for microbes and viruses. Now environmental samples can be measured by ‘omics - allowing observations of biological identity and function to be gathered at scope and detail never before available. As we learn about the diversity of microorganisms<sup>21</sup> and their associated biogeochemical processes, our view of the world’s ocean ecosystems has been expanded to appreciate the relevance of microbes<sup>22</sup> to all aspects of life, including ocean resiliency and marine resource management. ‘Omics is being used to conduct *Massively Parallel Sequencing at Planetary Scale*: Microbes, found throughout the oceans, soils, and animals, help control ecosystem services, including food production. Understanding how environmental conditions affect this critical component of the Earth system can help anticipate how communities will adapt as conditions change. In the fall of 2017, the Earth Microbiome Project (EMP) released the first reference database, or atlas, of microbes covering the planet. This new database catalogues the microbiomes of Earth, which can help launch innovations in aquaculture,

<sup>21</sup> <https://oceanexplorer.noaa.gov/about/what-we-do/media/marine-microbes-workshop-report.pdf>

<sup>22</sup> <https://www.aoml.noaa.gov/keynotes/PDF-Files/May-June2016.pdf>

agriculture, energy, medicine, and the built environment. The Nature paper describing the catalogue (doi:10.1038/nature24621) received NOAA's 2018 Paper of the Year Award in the Oceans and Great Lakes category, was tweeted

1245 times and picked up by 29 news outlets when first released, and was cited 119 times in less than a year<sup>23</sup>.

## NEWS & VIEWS

### Crowdsourcing Earth's microbes

A large-scale study has been assessing microbial diversity by analysing DNA sequences from samples submitted by scientists around the globe. The initial results are now being used to create an open-access resource.

#### *NOAA-CalCOFI Ocean Genomics (NCOG):*

The NOAA-CalCOFI Ocean Genomics

(NCOG)<sup>24</sup> project is one example of a NMFS-

OAR Joint Activity. NCOG is helping enhance an existing ecosystem observation program by adding technology for genetic sampling as part of the California Cooperative Oceanic Fisheries Investigations (CalCOFI) program. CalCOFI is a premier U.S. ocean monitoring program that has been collecting information on fish and oceanography off the west coast of the United States to support resource management decisions since 1951. The program seeks to gain a comprehensive understanding of the dynamics of the California Current ecosystem to foster stewardship, resilience, and sustainable resource management. By introducing these genetic sampling technologies, scientists hope to improve understanding of the diversity and functions of microbes in the ecosystem, enabling better prediction of ecosystem response to environmental pressures including climate change. Compared to traditional chlorophyll and zooplankton volume measurements, the new biodiversity data provided by DNA sequencing is information-rich because genera and species are now identified. High throughput laboratory and bioinformatics protocols streamline data throughput compared to methods such as pigment analysis or visual inspection. This project<sup>25</sup> is a collaboration between AOML, the J. Craig Venter Institute (JCVI), and the Scripps Institution of Oceanography (SIO) with support provided by labs and programs of the National Marine Fisheries Service (NMFS), NOAA's Oceanic and Atmospheric Research (OAR), the NOAA Office of Ocean and Exploration Research (OER), AOML, and the U.S. Integrated Ocean Observing System (U.S. IOOS). Data from the NCOG project being transitioned to ecosystem models through a joint project involving AOML, the Southwest Fisheries Science Center (SWFSC), and JCVI through support from the AOML 'Omics Program. Ecosystem models help keep track of complex interactions and non-linear responses to drivers; this project is working to integrate molecular biology and physical oceanography into ecosystem models to improve the models' prediction power and utility to resource and conservation managers.

*Bioinformatics Enhancement:* Genome-based techniques improve our ability to characterize and monitor ecosystems, including characterizing environments with potential commercial potential. Examples include genomic signatures that mark natural resources such as oil and gas reserves, enzymes with commercial potential such as pharmaceutical or bioremediation potential, microbiomes that control the health of commercially valuable wild and aquaculture species, and indicators of anthropogenic stress which can forecast degraded environmental

<sup>23</sup> <http://www.aoml.noaa.gov/keynotes/PDF-Files/Oct-Dec2017.pdf>

<sup>24</sup> <https://calcofi.org/field-work/bottle-sampling/ncog-project/525-noaa-calcofi-genomics-project.html>

<sup>25</sup> <https://www.aoml.noaa.gov/keynotes/PDF-Files/Nov-Dec%202014.pdf>

productivity. However, our ability to supply bioinformatics expertise has not kept pace with the generation of sequence data, creating a data and analytical backlog and hindering transition of data collected into actionable information. To address this gap, AOML has been working to develop bioinformatics capacity, which is important to the success of all ‘omics projects. AOML has secured servers dedicated to bioinformatics analysis, hired young scientists to help with analysis, created user groups (local and NOAA-wide) to provide support, provided protocols, established a GitHub portal<sup>26</sup>, developed a sequence analysis pipeline, and created coding tutorials. The purpose of the AOML GitHub page is to share code across the lab and with the public, especially code that might be reused by others. Anyone at AOML who develops any kind of code (MATLAB, R, Python, FORTRAN, etc.) for any kind of research (hurricanes, ocean chemistry, ecosystems, or physical oceanography) is encouraged to create repositories and host them on the AOML page. A course “Python for Data Analysis” provides an introduction to data science using the Python programming language. The course was taught at the Scripps Institution of Oceanography from October 1 to December 5, 2018. The lectures and lessons, respectively, are available online YouTube and GitHub<sup>27</sup>. To facilitate analysis of amplicon datasets, a bioinformatics pipeline (“Tourmaline”) was created. Tourmaline is a DNA sequence processing workflow for Illumina amplicon sequence data that uses QIIME 2 and the software packages it wraps. Tourmaline manages commands, inputs, and outputs using the Snakemake workflow management system, facilitating rapid and reproducible amplicon sequence analysis.

## **Pacific Marine Environmental Laboratory (PMEL) ‘Omics Laboratory**

The Genetics and Genomics Group (G3) at PMEL was initiated in the fall of 2016. The group has a large number of active projects that utilize metagenomics, defined as the simultaneous study of genetics and genomics of entire communities using multiple targeted markers and genes and high-throughput sequencing. The focus is to discern marine and aquatic community dynamics (in terms of space, time, and/or life history). Samples include organisms (whole or in part) collected from water (e.g., plankton) or sediment, as well as gut contents and tissue samples. Particular focus is identifying (to species, and selected taxa to population) and quantifying relative abundance of marine invertebrates and fishes, from coastal, offshore, hydrothermal vent, and methane seep communities in the northeastern Pacific and Alaskan waters. Recent G3 PMEL publications using this custom targeted metabarcoding high-throughput sequencing approach and analyses include Klymus et al. (2017), Stepien et al. (2019) and Marshall and Stepien (2019). Projects compare biological community composition and relative representation of taxa with oceanographic parameters, including inorganic carbon chemistry, dissolved oxygen, temperature, and nutrients, to evaluate the influence of acidification, hypoxia, warming, and nutrients on biological communities. Paired eDNA and plankton samples are analyzed to compare community composition, and the group has been developing methods to accurately quantify the eDNA in metabarcoding libraries through the use of internal standards and positive controls. Benefits shown for metagenomic and eDNA analysis include:



<sup>26</sup> <https://github.com/lukenoaa/tourmaline>

<sup>27</sup> <https://github.com/cuttlefishh/python-for-data-analysis>

- Increased sensitivity and accuracy, and lower cost, over conventional sampling.
- Address prior inability to identify important species and populations whose early life stages cannot be visually discerned.
- High throughput identification of multiple taxa, from many samples, all at once.
- Joint analyses of genetic identity, taxon discrimination, and relative abundance representation at the community, species, and population-levels through multiple targeted metabarcoding high-throughput sequencing assays.

## Great Lakes Environmental Laboratory (GLERL) ‘Omics Program

One of the most important applications of the ‘omics program at GLERL is the prediction of harmful algal bloom (HAB) dynamics, toxicity, and ecosystem impacts. The cyanobacterium *Microcystis* is the dominant HAB species in the Great Lakes. Understanding its dynamics requires the use of ‘omics along with “traditional” methods (e.g., microscopy, growth, physiology). *Microcystis* produces the toxin microcystin. This taxon consists of many genotypes with variable traits, including the production of microcystin and other noxious secondary compounds, colonial forms, nutrients that can be used, and resistance to predation including by invasive mussels and viruses. *Microcystis* types and traits vary depending on environmental conditions, and molecular identification is needed because evaluation by microscopy does not necessarily discern these traits. ‘Omics data are being gathered to provide answers to how different *Microcystis* genotypes evade invasive mussel grazing through production of microcystin and other traits, and how this is a function of nutrient conditions. Hence, this directly supports GLERL’s HABs forecasting goals targeting the biomass of HABs and their toxicity and the impact of invasive species.



In the future, the aim is to use ‘omics to evaluate the effects of a full range of environmental factors (nutrient loading, climate, other grazers, and viruses) in promoting HABs and their impacts on the ecosystem so that appropriate mitigation strategies can be developed. Additional efforts include work with NCCOS (see below), AOML, and the Monterey Bay Aquarium Research Institute (MBARI) to develop sensors on Environmental Sample Processors (ESPs) and autonomous vehicles to identify different strains of *Microcystis* and their potential toxicity. Another project at GLERL is in conjunction with PMEL (see above) to develop eDNA monitoring tools for planktonic communities, including early detection, identification, and genetic variation of invasive species at multiple taxonomic levels - from communities, through species, and their populations using targeted metabarcoding high throughput sequencing assays (see examples in Klymus et al. 2017, Stepien et al. 2019, Marshall and Stepien 2019).

## Ocean Acidification Program (OAP) Investment in ‘Omics

Ocean acidification research at NOAA is guided by the [2010 NOAA Ocean and Great Lakes Acidification Research Plan](#) and the [Ocean, Coastal, and Great Lakes Acidification Research Plan: 2020-2029](#), and led by the OAP. The OAP is housed in OAR, but serves as a matrix program

across OAR, NMFS, NOS, and NESDIS. The OAP funds work in NOAA labs and science centers via directed allocation and to non-Federal entities through competitive awards. To date, the OAP has funded NMFS (NWFSC and AFSC) to use metabolomic, genomics, and transcriptomics to better understand the sensitivity of Dungeness crab, red king crab, and coho salmon to ocean acidification with and without changes in other environmental conditions. Such research allows detection and quantification of sub-lethal impacts of ocean acidification. This type of information is vital for efforts to generalize understanding of species sensitivity to unstudied, related species and for efforts to scale-up from short-term laboratory studies to potential impacts on population dynamics in the wild.



Through a collaborative funding effort, the OAP and the Northeast Sea Grant Programs funded university researchers to use transcriptomics and genomics to better understand the sensitivity and resilience of three commercial shellfish species to ocean acidification and of American lobster to ocean acidification and warming. Information from these projects will inform wild-capture fishery management and aquaculture efforts in the region.

## **Office of Ocean Exploration and Research (OER) 'Omics Interest and Investment**

NOAA's Office of Ocean Exploration and Research (OER) is the only U.S. federal organization dedicated to ocean exploration. Our work results in the establishment of baseline observations in poorly understood areas of the ocean. Collected data are quickly made available to the science and management communities, delivering the environmental know-how needed to better understand deep-water (>200m) areas and more effectively target future exploration and research efforts. In addition to convening a 2011 NOAA Marine Microbes Workshop, supporting the MBON Demo Projects and NCOG Project, OER supports 'omics science and technology projects via its Federal Funding Opportunity (FFO) solicitations and Cooperative Institute agreements. OER's 2019 FFO solicited proposals focused on the discovery of microorganisms, sponges, corals, and other organisms with biopharmaceutical or biotechnical potential. Prior opportunities and agreements have enabled such things as:



## **Ocean Exploration and Research**

- the use of next-generation DNA sequencing technologies and metagenomic sequencing to probe the diversity, distribution, and functional roles of microbes in the Arctic
- investigation into the abundance of microbes in volcanic rocks of different ages as well as specifics about what microbes are present in the Gulf of Alaska; and study of the genetic structure of deep-sea gorgonian corals to determine whether seamount populations are genetically isolated units
- study of the various complex networks within Arctic and Antarctic sea ice that intersect to make up this unique, extreme ecosystem: brine channels, food webs, and genetic exchange networks that are all dominated by microbes

- exploration in the Gulf of Mexico for novel bioactive compounds from marine organisms that have potential as pharmaceutical products or biomedical research tools
- support of cross-line activities in ‘omics. ‘Omics was included as a topic in three bilateral (NMFS-OAR) meetings held during the years 2014-2016, for which a white paper and synopses were produced, and the NOAA Marine Microbe Workshop (mentioned briefly above and also sponsored through the NOAA Oceans and Human Health Initiative) was held at the Hollings Marine Laboratory in 2011.

## Sea Grant Investment in ‘Omics

The National Sea Grant College Program supports ‘omics projects that span the program’s four focus areas: Healthy Coastal Ecosystems, Sustainable Seafood and Aquaculture, Resilient Communities and Economies, and Environmental Literacy and Workforce Development. Sea Grant is committed to advancing the ‘omics field. For example, South Carolina Sea Grant researchers have developed a streamlined genetic analysis pipeline which allows for broader collaboration and distribution of large quantities of genomics data among current and potential investigators. Other ‘omics projects include:



### ***Healthy Coastal Ecosystems***

*Pathogen detection in marine mammals:* In 2015, Oregon Sea Grant supported an exploratory project using metagenomics on brain tissues of recently stranded and deceased harbor seals to assess whether undetermined causes of stranding resulted from novel viral infections. Deep sequencing technology, synthesized in partnership with Oregon State University’s Center for Genome Research and Biocomputing, captured viral diversity and was used to overcome RNA contamination from host cells. This technology is unique because it allows for high sequencing at a low cost. A novel bioinformatics pipeline was also developed to ensure accurate analysis.

*Conserving endangered species:* A Texas Sea Grant-funded project is currently conducting genetic analyses of Kemp’s ridley sea turtle populations in Texas and Mexico to determine the genetic variability of the species, which is the most endangered sea turtle in the world. The researchers have already discovered and optimized new microsatellite markers and discovered genomic protocols that can be used by other scientists researching Kemp’s ridley and potentially other sea turtle species.

*Analyzing responses of planktonic invertebrate communities to physical oceanographic parameters:* A Washington Sea Grant program in partnership with the G3 Lab at NOAA PMEL is developing high throughput metabarcoding sequencing analyses to understand responses of planktonic communities to oceanographic conditions, including warming, hypoxia, and acidification. Special focus is on copepods and bivalves (clams, mussels) at early life stages.

### ***Sustainable Seafood and Aquaculture***

*Genomic profiling to improve finfish aquaculture:* North Carolina Sea Grant-funded researchers provided critical research supporting a hybrid striped bass industry that now produces annual revenues in excess of \$10 million in North Carolina alone. The team's generation of large databases of miRNA sequence data aided the National Animal Genomics Project. The team's discovery of scores of novel genes is expected to lead to the determination of differences in gene expression between poor and good quality oocytes and eggs. Such new knowledge promises to enhance production success and profitability in aquaculture enterprises.

*Detecting Vibrio in the environment:* In 2017, a team of New Hampshire Sea Grant researchers applied a newly developed detection method for *Vibrio parahaemolyticus* to identify two strains of concern in environmental samples from three New England states. These non-hypervirulent strains proved useful in ongoing comparative genomics analysis of pathogen evolution and confirmed the absence of hypervirulent strains in NH coastal waters.

*Controlling Listeria in seafood:* Control of the foodborne pathogen *Listeria monocytogenes* is a top priority for the seafood industry and public health agencies and represents a considerable challenge for all food processors. New York Sea Grant has funded several projects led by researcher Dr. Martin Wiedmann at Cornell University spanning over a decade (from 2000 – 2014) examining various aspects of *L. monocytogenes* growth and control. This research specifically used genomics-based methods that led to the development of a manual incorporated into the Association of Food and Drug Officials' (AFDO) model "Cured, Salted, and Smoked Fish Establishments Good Manufacturing Practices (GMP)."

### ***Resilient Communities***

*Reducing harmful algal blooms (HABs):* Ohio Sea Grant-funded researchers are using metagenomics analysis to examine samples collected by the Ohio EPA during routine cyanobacterial screenings. They are focusing on genes that produce saxitoxins and cylindrospermopsis, another group of toxins of emerging concern in Ohio. The information will be used to help water treatment plants improve their water reservoir management strategies to prevent potentially toxic blooms from forming.

### ***Environmental Literacy/Workforce Development***

*Private sector partnerships advance human health:* Minnesota Sea Grant's funding of fish genetics research has led to revolutionary advances in human gene therapy. In 1997, Minnesota Sea Grant funded a project that aimed to engineer faster-growing fish for aquaculture purposes. This project led researchers to discover and transform an inactive gene originating in the salmon family into a non-viral DNA delivery system called the Sleeping Beauty (SB) transposon system. This system can transport normal genes into genetically damaged cells that cause such things as retinal degeneration and cancers in humans. The initial support from Sea Grant has since spurred private sector interest and more than \$5 million in direct funding to support further research in SB's application to human gene therapy. Since its creation, SB has inspired the founding of Discovery Genomics, Inc. by Minnesota Sea Grant researcher Perry Hackett, Ph.D.



and made significant advancements in human gene therapy through laboratory experiments on mice. More recently, researchers at the University of Texas MD Anderson Cancer Center used SB to perform clinical trials on human patients with blood disorders – including lymphocytic leukemia and non-Hodgkin lymphoma; no severe or long-term toxicity was found among any of the 33 patients, and 47% of patients treated remained in complete remission or showed disease regression.

## National Ocean Service (NOS) Activities in ‘Omics

### ‘Omics at the Hollings and Beaufort Marine Laboratories

The Hollings Marine Laboratory (HML) in Charleston, SC is part of the National Centers for Coastal Ocean Science (NCCOS) laboratories in NOAA’s National Ocean Service (NOS). HML is a unique collaboration that brings together basic, applied, and biomedical research expertise to provide support and biotechnology applications to sustain and protect healthy coastal ecosystems and identify linkages between environmental and human health. It is the product of a long-term Joint Project Agreement between NOAA, the National Institute of Standards and Technology, the South Carolina Department of Natural Resources, the College of Charleston, and the Medical University of South Carolina.



Since opening in 2001, scientists at the HML have been active in a variety of ‘omics disciplines including genomics, metagenomics, transcriptomics, proteomics and metabolomics. The laboratories and instrumentation support genomic and transcriptomic research. Mass spectrometers and nuclear magnetic resonance spectroscopy capabilities support metabolomics and proteomic applications (see Appendix 3).

Examples of ‘omics research at HML have included the following investigations:

- Restoration genomics – using genomics to inform coral restoration practices based on understanding adaptation and acclimation potential of source coral stocks for nursery propagation and siting coral outplant.
- Gene expression in non-model estuarine organisms using transcriptomic technologies to develop predictive and quantitative biosensors of ecosystem health. Work has been conducted on a variety of organisms (e.g., coral, fish, dinoflagellates) at different trophic levels in an effort to assess environmental health from a whole ecosystem perspective.
- Microbial source tracking and pathogen detection using metagenomics and transcriptional profiling. This work includes development of assays for zoonotic detection in marine mammals.
- eDNA analyses to examine microbial community change in response to environmental stressors (both natural and anthropogenic).

- Molecular level detection, identification, and quantification of harmful algal species, algal toxins, and toxin genes on autonomous moored and mobile (e.g., AUV) subsurface platforms in marine and freshwater systems.

Research at HML and its Cooperative Institute partners includes diverse topics such as wildlife epidemiology, health assessments for sentinel species, development of molecular diagnostic tools, harmful algal bloom mitigation, aquaculture, fisheries science, pollution chemistry, biotechnology development, and the valuation of human dimensions of coastal ecosystems and communities. The partnership structure of the joint project agreement at HML allows scientists to collectively address NOAA priority areas. The resulting synergy produces exciting and innovative approaches that address issues of immediate concern to coastal, marine, health and civic professionals. Furthermore, the NCCOS Information Technology group has enabled transitions to cloud-based high-performance computing analysis, and that experience could serve NOAA's 'omics enterprise.

Research is additionally underway at the NOAA Beaufort Laboratory in collaboration with the University of Maryland to investigate the transcriptomics of the organisms responsible for causing ciguatera fish poisoning the largest cause of nonbacterial seafood poisoning globally. The project is aimed at better distinguishing toxic from non-toxic species and identifying toxin production pathways that can be used as basis for qPCR assays capable of identifying regions at greatest risk for producing toxic shellfish and fish.

### **Coral Reef Conservation Program (CRCP) Interest and Investment in 'Omics**

Although CRCP technically sits in NOS, it is a true matrix program. It is interdisciplinary and collaborative across lines, traits required for success of NOAA 'omics activities. Coral reefs in U.S. jurisdictions are suffering from record levels of bleaching and disease. The potential for 'omics to mitigate the severe threats to corals worldwide, including bleaching and disease exacerbated by increasing sea surface temperatures, includes improving reef resilience through management of reef connectivity, studying stress, resiliency, and disease tolerance in different species of corals, and assisted adaptation. Assisted adaptation includes directed breeding, but genetic modification is considered due to the severity of threat. Management will apply research results to devise intervention strategies for population enhancement and recovery.



### **United States Integrated Observing System (U.S. IOOS) Investment in 'Omics**

The U.S. IOOS Program Office is housed in NOS. U.S. IOOS is a national-regional partnership working to provide new tools and forecasts to improve safety, enhance the economy, and protect our environment. Integrated ocean information is available in near real time, as well as retrospectively. Easier and better access to this information is improving our ability to understand and predict coastal events and impacts to marine life and human activities. The majority of IOOS appropriated funding is provided to external partners via cooperative agreements; the primary mechanism for distribution of these funds is through the IOOS Regional Associations

(RAs). NOAA has certified all 11 RAs as “Regional Information Coordination Entities” ensuring that data from all IOOS partners adheres to common federal collection, storage and management standards, meaning it can be integrated with other data, and help make "big data" research and development possible.



IOOS has invested in ‘omics activities for a number of years, providing funds for U.S. Marine Biodiversity Observation Network (U.S. MBON) testing and application of eDNA methods for operational observing, AOML coral reef metagenomics research and monitoring, eDNA comparison with ship-based sampling methods in Monterey Bay, kickoff of the NOAA-CalCOFI Ocean Genomics (NCOG) effort, and Environmental Sample Processor (ESP) deployments in multiple regions. In 2020, IOOS funded a series of workshops focused on eDNA methods and engagement with makers of ‘omics technologies (hosted by the Alliance for Coastal Technologies and developed with OAR/AOML and a non-Federal expert committee).

The IOOS Ocean Technology Transition program (~\$2M/year) sponsors the transition of emerging marine observing technologies to operations. Potential projects must show an existing operational requirement, demonstrate commitment to integrate into operations, and be within readiness levels 6 through 8 to qualify. In addition to serving as a testbed for deployments of new technologies on ocean observing platforms and evaluation of new sensors for operational settings, IOOS brings observing infrastructure, data management expertise and resources, and robust stakeholder engagement opportunities to the ‘omics enterprise.

## **United States Marine Biodiversity Observation Network (U.S. MBON)**

U.S. MBON is a key contribution to the biological component of U.S. IOOS. To preserve the ecosystem services upon which humans rely, long-term data about the distribution, abundance, habitat, and movement of multiple species (microbes to whales) is needed. ‘Omics analysis is one key to meeting the IOOS goal of providing operational, long-term biological observations and access to biological data. As such, IOOS is investing in, implementing, and managing multi-agency biological observing systems. All U.S. MBON projects (starting in 2013 and ongoing) incorporate some aspect of ‘omics, particularly with regard to advancement of eDNA technologies and integration of molecular data into operational streams. MBON partners are developing and testing eDNA methods to define the scope of potential applications, publishing best practices and methods as well as papers validating eDNA testing methods to identify organisms from microbes to marine mammals, applying MBON eDNA methods to improve screening for toxic algae at the State level (e.g., with the Florida Fish and Wildlife Conservation Commission), using eDNA metabarcoding to assess vertebrate diversity at oceanographic stations in Monterey Bay, and sampling coral spawning events to determine if eDNA can detect diversity of spawning corals, sponges, and brittle stars. Other U.S. MBON focal areas related to ‘omics include products to support the United Nations Sustainable Development Goal 14 (SDG14), which focuses on marine resources, and the United Nations Decade of Ocean Science for Sustainable Development (2021-2030).



MBON is implemented through the National Oceanographic Partnership Program (NOPP); U.S. IOOS, OAR/Office of Ocean Exploration and Research, NMFS, the National Science Foundation (NSF), NASA, the Office of Naval Research (ONR), and the Bureau of Ocean Energy Management (BOEM) have provided funding for MBON since its inception in FY2014. Shell Oil also provided initial investment for Arctic biodiversity monitoring. New projects were awarded in FY2019 through the NOPP BAA topic “Sustained observations of marine biodiversity for improved understanding of marine ecosystem responses to changing environmental conditions,” expanding geographic coverage of the Network with plans for further exploration of eDNA methods, applications, and feasibility for an operational observing system, as well as integration of eDNA with other observing approaches.

MBON supports integrative activities and development of new technologies and approaches to describe the status of marine life and ecosystems in the face of change and multiple stressors. US MBON is the national contribution to the global MBON and is intended as a long-term, multi-agency, multi-disciplinary and multi-sector effort.

## **NOAA Information Enterprises and ‘Omics**

Advances in medical science have enabled progress in environmental and other sciences. However, the production and analysis of modern DNA sequencing is an example of Big Data, which is why the tool is both transformational and disruptive. The demand for ‘omics information has driven huge computational demand, with biology rivaling the needs of astrophysics or the Internet (Stephens et al. 2015). The need for bioinformatics expertise and computing power is a priority identified across NOAA Line Offices and across federal agencies, in general (Stulberg et al. 2016, FTAC-MM 2015) because ‘omics analyses are generating an extreme demands in bioinformatics expertise, computational resources, data storage, and data management. Strategic partnerships are one approach NOAA has used to address gaps in these areas. However, feedback from the OTF survey revealed a need for internal support and concerns of an overdependence on extramural expertise and resources.

In 2018, the NOAA Office of the Chief Information Officer (OCIO) investigated mechanisms to procure cloud services to meet many of the ‘omics big data challenges, such as Cloud Utility Contract procurement. Aside from the scalable computational power provided by major cloud service providers (CSPs), e.g. IBM, Google, Amazon Web Services (AWS), and Microsoft, all offer specialized compute capabilities specifically for genomic processing. NOAA OCIO is seeking a multi-cloud, Blanket Purchase Agreement (BPA)-based solution that will allow NOAA scientists to choose the platform and tools that they need, on any given data, to accomplish their goals.

Additionally, advances in Machine Learning (ML) and Artificial Intelligence (AI) techniques show promise in interpreting results and ‘thinning’ input data streams. The NOAA OCIO sees the enormous potential in applying the combination of highly scalable compute resources and AI and ML frameworks to ‘omics big data. The demands of AI technologies, which includes instant

access to all of NOAA's stored/archived data as well as new data sources simultaneously, will significantly influence changes to NOAA's data architecture.

The National Centers for Environmental Information (NCEI) within the NOAA Line Office of National Environmental Satellite, Data, and Information Service (NESDIS), supports the NOAA's 'Omics Strategy by helping the 'omics data community develop and link metadata standards and 'omics sampling events to the NOAA environmental data community.

## NOAA External Engagement in 'Omics

As part of the portfolio analysis, *a brief and non-exhaustive* summary of external engagement in 'omics is summarized below.

*National Conference on Marine Environmental DNA:* The first national conference on marine eDNA was held November 29-30, 2018<sup>28</sup>. The goal was to help accelerate marine eDNA science and application by bringing together researchers, government agencies, and private foundations. Major themes of the conference included: **Technology Development:** Faster, cheaper, more portable; **Bioinformatics:** Genetic reference databases, analytic software, data compatibility; **eDNA Biology:** Relating eDNA presence and abundance to organism presence and abundance. The conference sought to foster a commitment to cooperative research efforts, such as standardized protocols, baseline sites for ongoing monitoring, technology development, and building genetic reference databases. The 2<sup>nd</sup> National Conference on Marine Environmental DNA is planned for 2021.

*U.S.-Norway Intergovernmental Group on eDNA Implementation for Fisheries Stock Assessments and Management (UNIG):* This international collaboration was initiated in 2018 to bring together bi-lateral expertise for identifying challenges and potential solutions associated with implementation of eDNA as a monitoring tool for knowledge-based ecosystem management and fisheries stock assessments (UNIG 2020). The goal of this consortium is to develop a strategy for future implementation of eDNA for fisheries management, leveraging resources and international expertise to address concerns of stakeholders. Research priorities and plans to achieve outcomes identified from round table discussions include: best practices for sampling, sample processing and data analysis; standardization of protocols; description of case studies where eDNA can generate/improve knowledge; execution of *ad hoc* pilot studies; performance of interlaboratory calibration studies; development of standard reference material; sharing of biological and genetic resources across national borders; initiation and execution of bi-lateral field activities testing eDNA implementation; effective communication of eDNA-based results.

*Atlantic Ocean Research Alliance (AORA) Marine Microbiome Group:* The [Marine Microbiome Roadmap](#) (AORA 2020) was launched within the AORA framework. The AORA collaboration be-

---

<sup>28</sup> [https://phe.rockefeller.edu/eDNAmarine2018/docs/MURU\\_eDNA\\_Conference\\_final\\_report.pdf](https://phe.rockefeller.edu/eDNAmarine2018/docs/MURU_eDNA_Conference_final_report.pdf)

tween Canada, the European Union, and the United States helps implement the Galway Statement on Atlantic Ocean Cooperation. The Roadmap was developed to help align and advance an Atlantic-wide approach to microbiome research and monitoring to (a) facilitate cooperation along and across the Atlantic, (b) increase the visibility of marine microbiome research, and (c) promote a healthy ocean and sustainable bioeconomy. The Roadmap identifies three thematic pillars (Environment and Climate, Food Value Chain, and Biodiscovery) supported by key cross-cutting elements to support a mission to better understand, protect, predict, and sustainably harness the marine microbiome.

*Global 'Omics Observation Network (GLOMICON):* This group was initialized in 2018 at an international meeting in Bremen, Germany (Peplies et al. 2018), which brought together researchers and practitioners who use multi-omic technologies to investigate the biology and ecology of the biosphere. The group collects protocols which involve or have been developed through collaboration by GLOMICON nodes or that have been tested across multiple institutions and designed to maximize consistency<sup>29</sup>. This group



The above list is not exhaustive national nor internationally. For example, a U.S.-Japan Living Resources Bilateral meeting was held in May 2018 to discuss eDNA and metagenomics. The Korean-U.S. Bilateral Agreement represents another potential avenue for international collaboration<sup>30</sup>. Other groups NOAA is active with includes the National Microbiome Data Collective (NMDC), the Genomics Standards Consortium (GSC), and Biodiversity Information Standards (TDWG) and associated groups.

## **NOAA Cross-Agency Engagement in 'Omics**

*Microbiome Interagency Working Group (MIWG):* The MIWG released an Interagency Strategic Plan for Microbiome Research (MIWG, 2018). Microbiome research focuses on understanding the structure, function, and interactions of microbial communities. The strategy, drafted through collaboration of 23 federal agencies, recommended three areas of focus: 1) Support interdisciplinary, collaborative research to enable a predictive understanding to enhance public health, improve food and environmental security, and grow new bioeconomy product areas. 2) Develop platform technologies to generate critical insights and improve access to and sharing of data collected across ecosystems. 3) Expand the workforce through educational opportunities, citizen science, and public engagement. The MIWG was organized under the Life Science Subcommittee (LSSC) of the National Science and Technology Council (NSTC) Committee on Science (CoS).

*Interagency Working Group on Biological Data Sharing (IWGBDS):* The priority of the IWGBDS is to improve the ability to extract knowledge from ever-increasing volumes of diverse types of

---

<sup>29</sup> <https://www.protocols.io/groups/glomicon>

<sup>30</sup> <https://www.fisheries.noaa.gov/feature-story/nefsc-milford-lab-south-korea-connection>

biological data in order to expedite translation of basic information to new applications and technologies; use computational analysis to drive innovation and decision-making; and protect personal privacy and confidentiality. The IWGBDS will identify robust mechanisms for sharing data and deliver a federal biological data sharing road map for 2020-2025 to the Biological Science Subcommittee, Committee on Science, National Science and Technology Council.

*Government eDNA Working Group (GEDWG):* GEDWG is a group of scientists from a number of organizations including USGS, USFWS, NOAA, USACE, DOD, USDA, and academia. The group hosted workshops in 2016, 2017, and 2019. The 2019 workshop emphasized aquatic DNA, metabarcoding, and bioinformatics pipelines and processing with a focus on quality assurance and control. A practical session on applied field projects focused interpretation and use of eDNA data for environmental management decisions. The first day was an optional computer based intensive focusing on the application of Occupancy models and the use of the R package EDNAOCCUPANCY (developed by USGS). Additionally, statistical analysis of metagenetic/metabarcoding methods using a variety of software programs will be explored. This meeting was immediately followed by a bilateral meeting on eDNA between the USGS and NOAA.

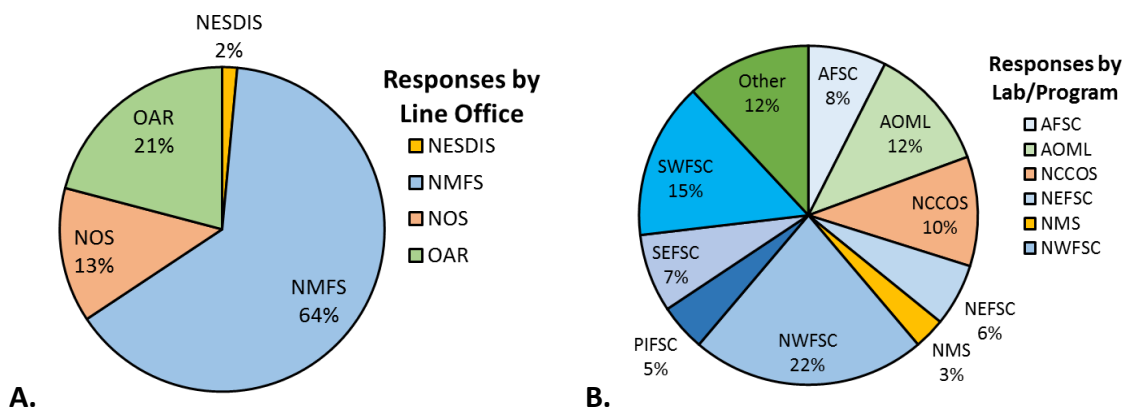
# APPENDIX II.

## 'OMICS TASK FORCE SURVEY RESULTS AND INFRASTRUCTURE REVIEW

### Survey Results

To inform the NOAA 'Omics Strategy, the OTF launched in 2018 a broad survey targeting current and potential users of 'omics research and technologies to better understand the breadth and scope of 'omics activities pursued across NOAA. The goal of the survey was to broadly quantify activities, to enumerate the types of uses and users of 'omics, and to identify major challenges, roadblocks and prospects impacting the realistic implementation of 'omics tools across NOAA.

The survey was broadly disseminated through the NOAA Research Council membership and the OTF. The OTF compiled 67 responses. The response breakout by Line Office was 64% from NMFS, 21% from OAR, 13% from NOS and 2% from NESDIS (Fig. A1). The NWS and OMAO did not provide responses as this type of research is not part of their operational requirements.

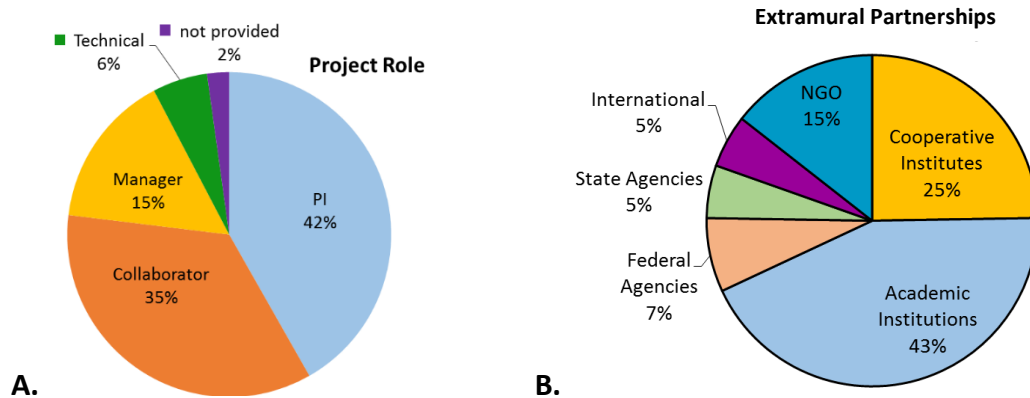


**Figure A1.** Breakout of survey responses by A) NOAA Line Office and B) by lab or program (other = 1.5% each; OAR: CPO, GLERL, NSL, OER, PMEL, Sea Grant; NESDIS: STAR). See List of Abbreviations.

The majority of survey respondents reported that they currently use some type of 'omics technology in their research (87%, n=67). Most survey respondents (88%) plan to use 'omics technologies in future research and 81% would be open utilizing NOAA expertise in the future. 'Omics technologies in regard to the survey were defined as next generation biotechnologies which have the capacity to generate large volumes of data, with examples including but not limited to genomic, transcriptomic, and metagenomic studies and could a variety of applications such population genetics, bacterial source tracking, and eDNA. The 'omics users were mostly in the roles of a project principal investigator, collaborator, or program manager, and they reported

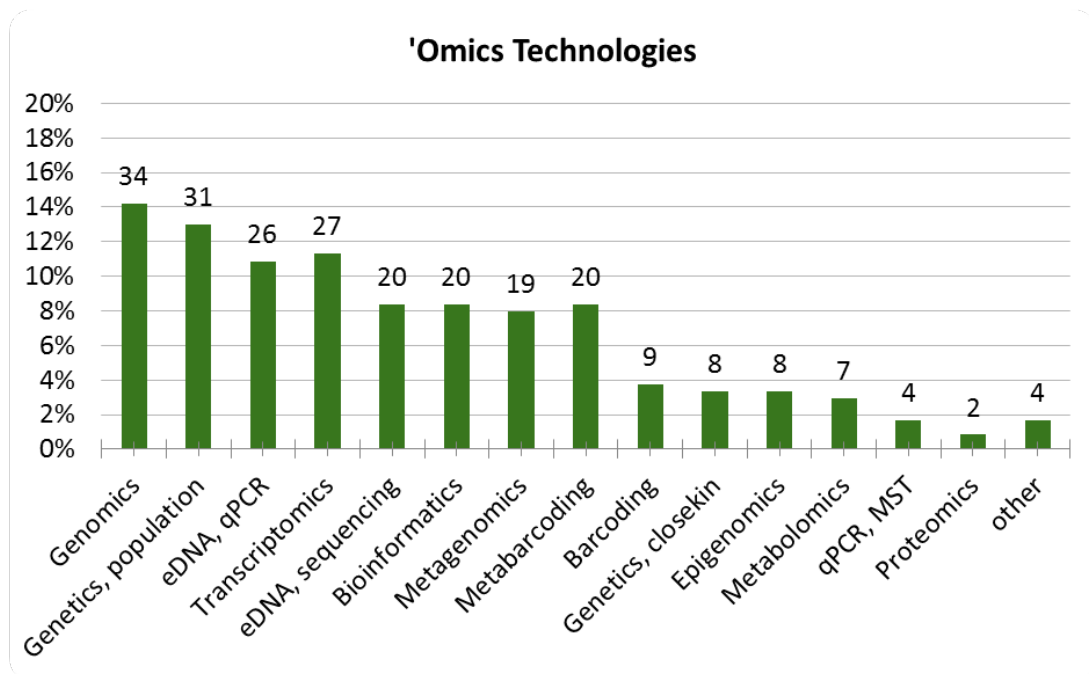


that they generally collaborated with academic institutions and NOAA Cooperative Institutes (Fig. A2), showing a deep engagement of NOAA 'omics science with the extramural .

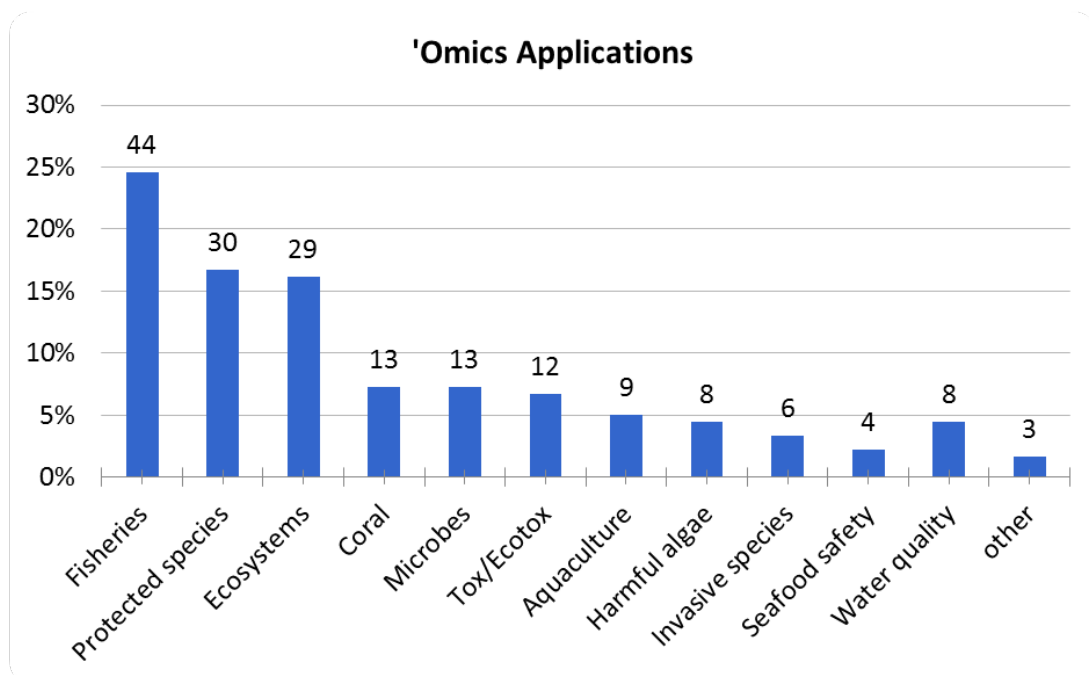


**Figure A2.** Project roles and partnerships A) based on responses to the survey question “Identify the capacity through which you use ‘omics technology” for any of the provided list of projects (n=91). (PI = Principal Investigator; Manager = program or project manager; science advisor/funder, Technical = technician, contractors, post-doctoral associates), and B) reported Extramural Partnerships (n = 91).

The most common types of ‘omics technologies reported to be in use were genomics (genome sequencing), population genetics, and eDNA analysis using qPCR. This was closely followed by transcriptomics, eDNA analysis using amplicon sequencing, bioinformatics, metagenomics, and metabarcoding (Fig. A3). The next group of reported applications included barcoding, genetics for stock assessment applications using close-kin mark-recapture, and qPCR for water quality applications (Microbial Source Tracking, MST). Research efforts using ‘omics approaches that are less commonly available, such as epigenomics, metabolomics, and proteomics were noted, suggesting these as emerging approaches to address research needs. Respondents reported that these ‘omics techniques were mostly used in fisheries research and surveys, protected species research, and ecosystem studies (Fig. A4), which is consistent with the distribution of survey respondents (Fig. A1). These applications were followed by research involving corals, microbes, toxicology or ecotoxicology, aquaculture, harmful algae, invasive species, seafood safety, and drinking and recreational water quality. Overall, these findings are consistent with the fact that NMFS, OAR and NOS provided the majority of responses, and the primary mission for these Line Offices is to support ecosystem level research and the development of living resource management products.



**Figure A3.** Types of investigative techniques used based on responses to the question “What types of ‘omics technologies do you use?” (n=239 responses, with values per category shown over each bar of the graph).



**Figure A4.** Types of ‘omics applications based on the question “How do you apply your ‘omics research?” (n=179 responses, with values per category shown over each bar of the graph).

One of the survey questions focused on challenges associated with ‘omics data. A lack of ability to analyze ‘omics data due to personnel and information technology limitations was cited as the primary impediment to implementing ‘omics approaches in NOAA. *Overall, impediments fell into two major categories - lack of expertise and IT capacity, with 51% and 46% of responses falling into these groups, respectively (Fig. A5).* Survey results indicated major gaps in the areas of staffing, computing capacity, and bioinformatics expertise. These results were generally consistent with a cross-agency portfolio analysis of microbiome science conducted in 2015 (MIWG, 2018). Out of 135 total responses, cited impediments were lack of:

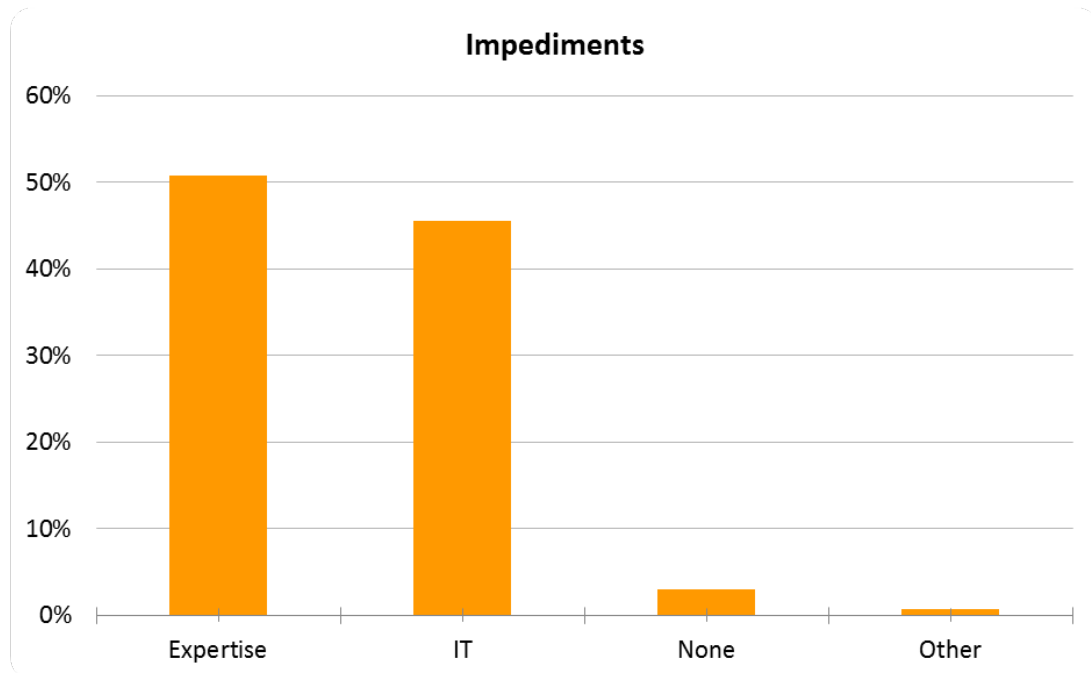
- Data Analysis Capabilities (“I don't have the time to do it myself or the staff to get it done”) – 36%.
- Data Processing Power (“I need more high-performance computing power”) – 24%.
- Data Analysis Expertise (“I have the data but I'm not exactly sure what to do with it”) – 18%.
- Data Analysis Resources (“I know what to do but I don't have access to the proper software/programs to analyze my data”) – 13%.
- Data Storage Capacity (“I'm analyzing, but I'm running out of space for all my data”) – 13%.
- Data Processing/Network Speed (“I have a bottleneck in uploading/ downloading /transferring my data”) – 9%.
- Data Acquisition (“I have an idea on what I would like to do but I don't know exactly what data I need or how to get it”) – 6%.

In addition to these responses, survey respondents were able to provide write-in options.

***Write-in responses describing dependence on extramural expertise as an impediment to implementation accounted for 6% of the total responses.*** Compelling write-in responses on this issue were received from NOS, NMFS, and OAR respondents. Problems gaining remote access to bioinformatic servers was also reported in the write-in responses (2%). Another issue raised was that science managers lack the expertise needed to prioritize funding or to evaluate projects (2%). Although this issue could also be classified as expertise, it was left as a separate category (“other” in Figure A5) because it is a distinct challenge. Specific feedback in the responses about impediments included:

- In multiple cases, lack of bioinformatic expertise was cited as a hurdle for being able to perform world-class research. Lack of internal analytical capacity was cited as a cause for conservation questions being addressed in a piecemeal fashion or not at all.
- Multiple respondents mentioned that limitations in computing and storage capacity resulted in data being housed with Cooperative Institute or academic partners.
- Any investment in HPC resources must include investment in systems administration (IT personnel familiar with research computing) and educational resources for researchers. Positive feedback for cloud services is that it minimizes the need for IT personnel, but there is no in-person help. Complaints about remote access to servers at NOAA facilities included: slow, tedious, and not allowed to do so (for a Mac server).

- Complaints of not being allowed to connect to cloud-based services from behind the NOAA firewall, including the use of applications offered in conjunction with in-house sequencers.



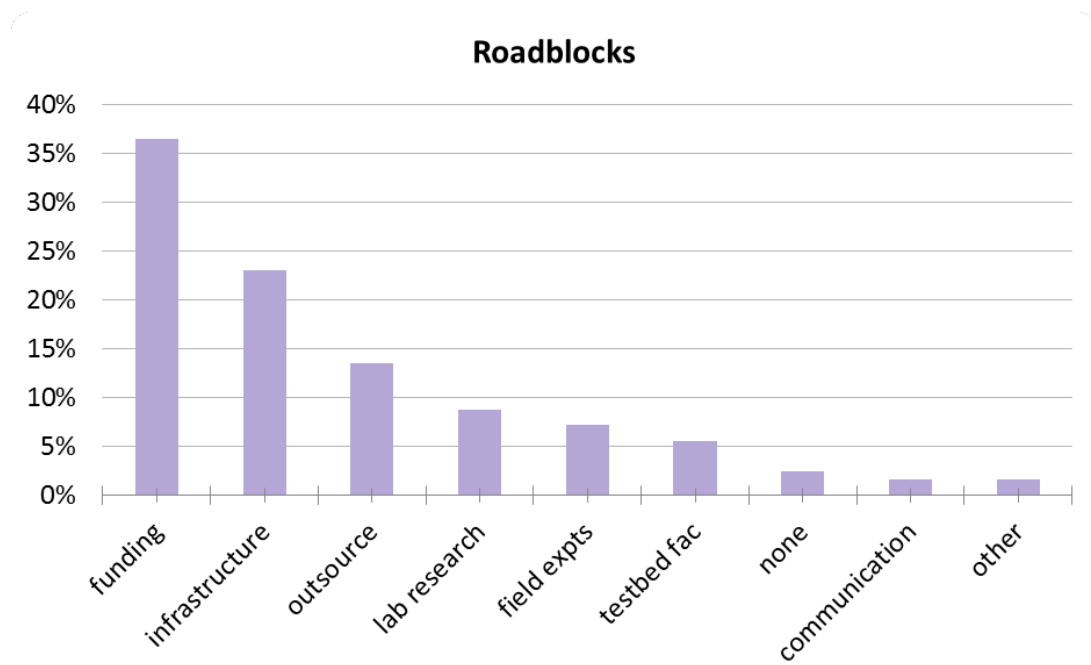
**Figure A5.** Breakout of survey responses (n=135) by impediments focused on analysis of ‘omic data. Expertise – data analysis capabilities (time/staff), data collection expertise, data analysis expertise, and dependence on extramural sources for expertise; IT – data processing power, data analysis resources (software), data storage capacity, data processing/network speed, remote log-in issues. Other – science managers lack expertise needed for funding.

Currently, the amount of ‘omics data produced by NOAA researchers appears manageable on an *ad hoc* basis. However, as ‘omics tools become commonplace across the agency, there is a real concern by the respondents for how NOAA will support the data storage and data processing requirements borne by ‘omics tools and technologies (see Table A2). ***These results suggest that sequence analysis is a choke point for transitioning data into actionable information. Furthermore, funding issues were consistently mentioned as a roadblock (Fig. A6), suggesting that researchers do not have a path forward to address these gaps.***

In addition to the question about impediments, which were data focused, one survey question looked at more general challenges facing implementation of ‘omics in NOAA. Two major roadblocks for ‘omics implementation cited by respondents were:

- lack of dedicated sources of ‘omics funding provided with appropriate timing and a duration (multi-year) consistent with research questions and development timelines
- lack of a strong, deliberate, and mission-driven commitment to an ‘omics infrastructure

Out of 126 total responses, major roadblocks to ‘omics implementation were related to issues of: funding (37%), operational infrastructure (23%), outsourcing/contracting (17%), laboratory research (11%), field experiments (9%), and test bed facilities (7%). Write-in responses included the need for better communication across NOAA to leverage expertise (2%), and science management concerns (2%, “other”) (Fig. A6). Several comments focused on the need to fund continued data collection, as well as data processing. One comment mentioned the need for longer-duration funding to properly study slow growing species, such as corals. Concerns over infrastructure included both sequencing and computing capacity. Based on details provided by respondents, it was clear that categories were not mutually exclusive. For example, the outsourcing/contracting bin garnered concerns about NOAA being over-dependent on external expertise and also comments about contract cycles being out-of-phase from field deployment needs, ‘omics funding arriving at labs after procurement deadlines have passed, and dependence on reimbursable agreements to do work. Specific feedback in the responses related to roadblocks included: a lack of clear funding lines to support the various stages of sampling, processing, and analyzing ‘omics projects; in-house capacity (personnel & equipment) often inadequate to support medium to large projects; spotty access to High Performance Computing. Overall, these issues represent challenges that must be overcome to accelerate the state of the ‘omics science at NOAA and to support the transition of research into applications and operations.



**Figure A6.** Breakout of survey responses (n=126) regarding general roadblocks to implementing ‘omics approaches within NOAA. Infrastructure – operational infrastructure; outsource – outsourcing and contracting (includes dependence on extramural expertise as well as issues with process); field expts – field experiments; testbed fac – testbed facilities.

Regardless of impediments and roadblocks, the majority of survey respondents provided a positive outlook for their prospective use of ‘omics technologies in the future (88%), and their willingness to utilize expertise within NOAA to move ‘omics research forward (80%), signaling their commitment to use ‘omics research and technologies to tackle current and emerging challenges, or to improve NOAA’s ability to deliver operational products and services to support resource management and stewardship requirements.

### Review of NOAA Infrastructure to Support ‘Omics

As part of the portfolio analysis, a survey of NOAA molecular biological infrastructure was conducted with results provided in Table A1. Information was not provided for all laboratories; therefore, additional capacities may be available in the agency.

**Table A1. Infrastructure Capacity Reported for NOAA Laboratories.**

Infrastructure	Lab	Number
Environmental Sample Processor, 2nd Generation (buoy)	NWFSC	2
	GLERL	3
Environmental Sample Processor, 3rd Generation (mobile)		0
HiSeq Sequencer (used for shotgun sequencing)		0
MiSeq Sequencer	AFSC, TSMRI	1
	NWFSC	2
	SWFSC, Santa Cruz	1
	SWFSC, La Jolla	1
Minion Sequencer	NWFSC	2
Sanger Sequencer	NWFSC	3
	SWFSC	2
	HML	1
	CCEHBR	1
digital PCR instruments (for eDNA applications)		0
qPCR instruments	AOML	1
	GLERL	1
	CCEHBR	1
	HML	2
	NWFSC	3
	SWFSC	3
Sample preparation robotics	HML	1
	NWFSC	3
	SWFSC	1

### Current Bioinformatics Infrastructure

There are few “true” high performance computing systems dedicated to bioinformatics in NOAA. Currently, the bioinformatics computer clusters at the NOAA Fisheries Northwest Fisheries Science Center provides an example of the current and growing gap in this area. The older existing cluster at the NWFSC, obtained in 2011, was initially built to provide HPC for low-memory parallel processes used to annotate genomes and other data from high throughput

'omics technologies. Over time, the system has been upgraded to include faster read-write speeds from the head node to compute nodes, additional high memory (RAM), high storage capacity compute nodes used for intensive processes such as assembly of genomes and the transcriptomes (the genes expressed in given tissues and individuals); that older system proved the utility of a locally managed resource for bioinformatics across projects, and led to the purchase of a second bioinformatics computing cluster through the NMFS Genomics Strategic Initiative (SI). This newer bioinformatics cluster, Sedna, was installed in late 2019. Sedna is a state-of-the-art computational resource for use across SI projects, and an even more recent investment in upgrades to the system will allow the addition of more compute nodes. Sedna consists of a head node (login node), NFS storage (63 Tb, RAID6), and fast scratch storage (43 Tb, RAID0) attached to 28 standard compute nodes and 3 high memory nodes (all with CentOS Linux 7.6.1810). Networking includes 100/1000/10 Gb and 1:1 EDR Infiniband. Compute jobs are executed using the SLURM (v18.08.8) scheduler. Bioinformatic software on the system includes that which is needed for genome and transcriptome assembly, genome alignments, BLAST sequence database search software, and software used to identify and call SNPs in sequence data. Other programs include software for population genetics (STRUCTURE), R, and other basic utilities. Sedna is managed by a small number of staff at the NWFSC, and the current challenge will be continued funding of the contract bioinformatics staff on the project. This position serves the critical function of maintaining and updating ever-changing bioinformatics software and facilitating the use of the resource through response to user inquiries and troubleshooting the use of the job scheduler.

### **Projected Needs**

Estimates of required capacities necessary to support NOAA's 'omics activities (see Table A2) are on a par with or more than NOAA's capacities required to support NOAA satellite operations, and NOAA must determine a strategy to meet this challenge. These rough estimates are based on information from a 2017 High Performance Computing Cluster (HPCC) proposal. See Appendix II for a more detailed discussion on computing capacity needs.

The OTF Survey results indicated a need to get additional information about 'omics-specific computing requirements. The following information was provided courtesy of the Knight lab, University of California, San Diego:

- Processing makes use of multiple cores. However, most bioinformatic tools do not parallelize. As such they are node locked, and this requires a larger footprint system to handle needs. Depending on the pipeline, this can push needs to 1TB+ of RAM. Traditionally, 1TB nodes were sufficient; however, it is increasingly common for jobs to need upwards of this. Nonetheless, jobs are often run on 256GB/32 core systems. Jobs can occupy nodes with 64 cores for days, if not weeks. Better parallelization has been seen in a couple applications, but they are few.
- The "best fit" for memory footprint seems to be >8GB/core. Most jobs currently run at this level and remain economical to build.

- Low latency networks are not required but are always useful. A large network pipe and fast disk are useful.
- Parallel file systems, if used, can be an issue due to small intermediate file sizes. Depending on what filesystems are used there are ways around this but it can impact other jobs' write speeds.

*Opinions on cloud services at the time of this writing:*

- Good for people with no infrastructure expertise in house.
- They are expensive for 'omics processing due to several factors.
  - Size of datasets being processed and the resulting output sizes.
  - Footprint of the systems needed to be stood up for processing.
- Data transfer speed is unknown but has potential to be a bottleneck.
- Data visualization tends to be expensive to do in the cloud, so cloud bioinformatic computations require a lot of back-and-forth that is not convenient compared to a local server.
- For the money that some labs spend on cloud services, it would be possible to stand up a hefty system that would meet the needs of 'omics and any other processing.

**Table A2. Estimated capacities required to support NOAA 'omics activities.**

Compute Need	2018, per lab	2022, per lab (est, assuming exponential data growth)	Proposed Mechanism
Storage	60 TB	1 PB	new Cloud Utility Contract (Q3FY19)
Transfer (Network)	1-10 Gbs	100 Gbs	NWave, with Direct connects to Cloud
Processing	100 Cores	3200	new Cloud Utility Contract (Q3FY19)
Architecture	Stovepiped	Optimized cloud platform, common shared services as well as in-house capacity	NOAALink/NMITS acquisition vehicles, HPCC funding
Staffing	1 Federal FTEs, 1 Contractors/Cooperators	2 Federal FTEs, 3 Contractors/Cooperators	SciTech, ProTech, NOAALink/NMITS



# APPENDIX III.

## SUMMARY OF TERMS AND GLOSSARY

A number of interdisciplinary fields utilize molecular biology to characterize the structure, function, and dynamics of an organism or organisms. The result is a plethora of terms that can differ among fields, between researchers, and change over time (e.g., the usage of “next generation”). In general, terms differ according to whether they refer to genetic, protein, or metabolic data. Another important categorization is whether biological material was obtained from individuals versus a mixed community (“meta”). A simplified glossary of ‘omics terms relevant to this document is provided below.

- **Bioprospecting** - the search for economically valuable genetic and biochemical resources from natural environments and habitats.
- **DNA sequencing** - Any method or technology that determines the order of the four chemical bases -- adenine (A), guanine (G), cytosine (C), thymine (T) -- that comprise a strand of DNA (deoxyribonucleic acid), and thus reveal the genetic code of an organism. High-throughput or massively parallel technologies produce massive amounts of sequence data per run.
- **Environmental DNA (eDNA)** - DNA collected from environmental samples. This “community DNA” is derived from multiple species rather than an individual. Initially used to study microbial communities, it is increasingly being used to gather information about higher trophic levels using DNA from sloughed or excreted cells.
- **Epigenetics** - the study of compounds and proteins that interact with DNA to alter gene expression. Methylation and histone modification, for example, modify DNA to turn genes on and off. Epigenomics is the study of DNA modification extended to whole genomes, and it can be used to investigate how the environment or disease can impact gene expression.
- **Genetics** - the study of genes and their involvement in heredity. Genes provide the instructions used to make functional molecules, such as proteins.
- **Genomics** - the study of an organism’s entire set of genes (the genome) and the structure, function, evolution, and relatedness of genomes.
- **Metabarcoding** - the extension of barcoding to provide information on community biodiversity. Metabarcoding extracts DNA from a community or organisms instead of from the tissue of an individual. Amplified and sequenced DNA from tissue is matched to a set of barcodes, which are short pieces of DNA (genetic marker) that identify an organism.
- **Metabolomics** - the large-scale study of small molecules produced during metabolism. A variety of chromatography and spectroscopy techniques are used to measure the biochemicals produced by cells or tissues to provide a snapshot or organism state.
- **Metagenomics** - the extension of DNA sequencing to the analysis of communities by using DNA recovered from environmental samples. Targeted sequencing of amplified DNA is typically employed to identify organisms whereas “shotgun” approaches are used to provide information on functional genes.

- **Metatranscriptomics** - an extension of transcriptomics to provide a profile of community function and activity. It involves the study of RNA sequences recovered from environmental samples, which contain a multitude of species.
- **Microbiome** - multi-species community of microbes in a specific environment. The ocean microbiome is described as an assemblage of bacteria, archaea, microeukaryotes (e.g., fungi, diatoms, dinoflagellates, and other protists), viruses, and mobile genetic elements that is the heart of marine food webs.
- **Molecular Source Tracking** - a qPCR application that targets specific organisms of the gut microbiome to determine the source of fecal contamination in waters (e.g., bird, dog, cow, human). The information is used to inform mitigation strategies and to evaluate the efficacy of management action.
- **'Omics** - a term for a set of genome-based technologies used to examine DNA, RNA, proteins, or small molecules from a variety of sample types ranging from single cells to communities to identify the organisms present, their metabolic status, and how they might be affected by changing conditions.
- **Proteomics** - the study of an organism's total proteins, which are the biomolecules that perform a vast array of functions within an organism. As opposed to sequencing technologies, mass spectrometry is used to analyze proteins.
- **qPCR** - Quantitative polymerase chain reaction is a laboratory method that quantifies the amount of target DNA contained in a sample.
- **RNA sequencing** - a method used to provide transcriptomic information primarily through the sequencing of mRNA (messenger ribonucleic acid), an intermediary molecule that creates proteins based on the instructions of DNA. RNA is less stable than DNA but more directly reflects cellular behavior by indicating which processes are active/inactive at a given time.
- **Synthetic biology** - a field of study that uses aspects of genetic modification to create artificial genes and new lifeforms.
- **Transcriptomics** - the study of using RNA sequencing to measure the transcriptome (the combined sum of an organism's RNA transcripts).

# APPENDIX IV.

## CONTRIBUTORS

The NOAA ‘Omics Task Force (OTF) is a cross-NOAA interdisciplinary team established by the NOAA Research Council in 2018 with its term extended in 2019. The OTF was tasked to document the portfolio of ‘omics research and activities across the organization, identify priorities, foster integration and communication, and formulate solutions to implementation challenges in this nascent sector of research. The OTF membership represented a variety of line offices and programs. OTF members from April 2018-July 2020 are detailed below.

<b>Organization</b>	<b>Members</b>
<b>OAR</b>	Kelly Goodwin, Co-chair & ‘Omics Strategy Executive Team
<b>NMFS</b>	Mark Strom, Co-chair & ‘Omics Strategy Executive Team
<b>NESDIS</b>	Kirsten Larsen
<b>NESDIS</b>	Scott Cross
<b>NMFS</b>	Cisco Werner, ‘Omics Strategy Executive Team Lead
<b>NMFS - Office of Policy</b>	Jeanette Davis, ‘Omics Strategy Executive Team
<b>NMFS - Big Data Project</b>	Jonathan O’Neil
<b>NMFS</b>	Krista Nichols
<b>NOS</b>	Thomas Greig
<b>NOS-NCCOS</b>	Felipe Arzayus
<b>NOS-IOOS, MBON</b>	Gabrielle Canonico
<b>NOS-CRCP</b>	Jennifer Koss
<b>OAR-OER</b>	Margot Bohan
<b>OAR-’Omics</b>	Katharine Egan
<b>OAR-OAP</b>	Shallin Busch
<b>OAR-Sea Grant</b>	Rebecca Certner, Vice Chair
<b>OCIO</b>	Ed Kearns
<b>OCIO</b>	David Layton
<b>TPO</b>	Derek Parks

Abbreviations: Coral Reef Conservation Program (CRPC), Integrated Ocean Observing System (IOOS), Marine Biodiversity Observation Network (MBON), National Centers for Coastal Ocean Science (NCCOS), National Environmental Satellite, Data, and Information Service (NESDIS), National Marine Fisheries Service (NMFS), National Ocean Service (NOS), Ocean Acidification Program (OAP), Oceanic and Atmospheric Research (OAR), Office of the Chief Information Officer (OCIO), Office of Ocean Exploration and Research (OER), Technology Program Office (TPO)

In addition to the OTF members, several people provided direct input to this document prior to submission to the NOAA Research Council. This input represented the following NOAA laboratories or programs: AFSC, AOML, GLERL, NCCOS, NEFSC, NGI, NWFSC, OCIO, PMEL, Sea Grant, SEFSC, SWFSC. The OTF is grateful for the input from the following individuals:

- Barbara Ambrose, Mississippi State University (logo creation)
- Eric Archer, Southwest Fisheries Science Center
- Greg Doucette, National Centers for Coastal Ocean Science, Stressor Detection and Impacts Divisions
- Ian Enochs, Atlantic Oceanographic & Meteorological Laboratory
- Jeff Guyon, Alaska Fisheries Science Center
- John Hyde, Southwest Fisheries Science Center
- Ed Kearns, Office of the Chief Information Officer, Environmental Data Management Committee
- Matthew Laretta, Southeast Fisheries Science Center
- Wayne Litaker, National Centers for Coastal Ocean Science
- Derek Manzello, Atlantic Oceanographic & Meteorological Laboratory
- Phillip Morin, Southwest Fisheries Science Center
- Thomas Noji, Northeast Fisheries Science Center
- Steve Ruberg, Great Lakes Environmental Research Laboratory
- Carol Stepien, Pacific Marine Environmental Laboratory
- Luke Thompson, Northern Gulf Institute, and Atlantic Oceanographic & Meteorological Laboratory (logo creation).
- Russ Vetter, Southwest Fisheries Science Center
- Hank Vanderploeg, Great Lakes Environmental Research Laboratory
- Cheryl Woodley, National Centers for Coastal Ocean Science

# APPENDIX V.

## REFERENCES

- Appeltans, W. et al. The magnitude of global marine species diversity. *Current Biology* 22(23):2189-2202 (2012).
- Bergman, P.S. et al. Detection of adult green sturgeon using environmental DNA analysis. *PLOS One* 11(4): e0153500 (2016).
- Bohmann, K. et al. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* 29(6): 358-367 (2014).
- H. Bolhuis, P.L. Buttigieg, K. GOODWIN, R. Groben; D. Ludicone, A. Lacoursière-Roussel, S. Pesant, S. Robinson, S. Björnsson, L.S. Erlendsson, M. Rae. The Atlantic Ocean Research Alliance - Marine Microbiome Roadmap. doi: 10.5281/zenodo.3632526 (2020).
- Bourlat, S.J. et al. (2013) Genomics in marine monitoring: new opportunities for assessing marine health status. *Marine Pollution Bulletin* 74:19-31 (2013).
- Erickson, R.A., et al. Detecting the movement and spawning activity of bigheaded carps with environmental DNA. *Molecular Ecology Resources* 16(4): 957-965 (2016).
- FTAC-MM. Report of the fast track action committee on mapping the microbiome. Product of the Life Sciences Subcommittee of the National Science and Technology Council. [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/ftac-mm\\_report\\_final\\_112015\\_0.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/ftac-mm_report_final_112015_0.pdf) (2016).
- Goodwin, K.D., et al. DNA sequencing as a tool to monitor marine ecological status. *Frontiers in Marine Science* 4(107): 1-14 (2017).
- Hansen, B.K., Bekkevold, D., Clausen, L.W., Nielsen, E.E. The skeptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries* 19(5): 751-768 (2018).
- Hering, D. et al. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research* 138: 192-205 (2018).
- Hunter-Cevera, J. et al. Review of the Department of Energy's Genomics: GTL Program. The National Academies of Science. <http://www.nap.edu/catalog/11581.html> (2006).
- Klymus, K.E, Marshall, N.T, Stepien, C.A. Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLOS One* 12(5): e0177643 (2017).
- Lacoursière-Roussel, A. et al. Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology* 53: 1148-1157 (2016).
- Lodge, D.M., Avery, S., Fluharty, D.L., Jackson, J. Potential impact of 'omics and other emerging genetic technologies on NOAA's mission. [https://sab.noaa.gov/sites/SAB/Meetings/2016%20Documents/August%202016%20Documents/Potential%20impact%20on%20NOAA%20of%20emerging%20genetic%20technologies%20Lodge\\_Final\\_v2.pdf](https://sab.noaa.gov/sites/SAB/Meetings/2016%20Documents/August%202016%20Documents/Potential%20impact%20on%20NOAA%20of%20emerging%20genetic%20technologies%20Lodge_Final_v2.pdf) (2016).
- Marshall, N.T., Stepien C.A. Invasion genetics from thousands of larvae and eDNA of zebra and quagga mussels using targeted metabarcode High-Throughput Sequencing. *Ecology and Evolution*. 00:1-24 (2019).

- Matlock G., Hilderbrand D., Gaynor J.E., Scheurer D., Valette-Silver N., Brown C., Green D., Phinney J., Turner E., Vo Dinh T. White Paper: No Transitions? No Outcomes! NOAA Technical Memorandum OAR PPE-2. [ftp://ftp.library.noaa.gov/noaa\\_documents.lib/OAR/PPE/TM\\_PPE/PPE-2.pdf](ftp://ftp.library.noaa.gov/noaa_documents.lib/OAR/PPE/TM_PPE/PPE-2.pdf) (2013).
- Microbiome Interagency Working Group. Interagency Strategic Plan for Microbiome Research FY2018-2022. April 2018. Microbiome Interagency Working Group (MIWG), an interagency working group under the Life Science Subcommittee (LSSC) of the National Science and Technology Council (NSTC) Committee on Science (CoS). [https://science.energy.gov/~media/ber/pdf/workshop%20reports/Interagency\\_Microbiome\\_Strategic\\_Plan\\_FY2018-2022.pdf](https://science.energy.gov/~media/ber/pdf/workshop%20reports/Interagency_Microbiome_Strategic_Plan_FY2018-2022.pdf) (2018).
- NAO. NOAA Administrative Order 216-105B: Policy on Research and Development Transitions. [https://www.corporateservices.noaa.gov/ames/administrative\\_orders/chapter\\_216/216-105B.html](https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_216/216-105B.html). (2016).
- National Marine Sanctuaries. Coral reefs support jobs, tourism, and fisheries. [florida-keys.noaa.gov/corals/economy.html](http://florida-keys.noaa.gov/corals/economy.html). accessed: December 2018.
- NMFS OST. Fisheries of the United States 2016. Current Fishery Statistics No. 2016 (2017).
- OSET ('Omics Strategy Executive Team). NOAA 'Omics Strategy. <https://nrc.noaa.gov/LinkClick.aspx?fileticket=RReWVFNjr5I%3D&tabid=92&portalid=0> (2020).
- Peplies, J., Janssen, F., and Buttigieg, P.L. Workshop in enhancing interoperability and coordination of long-term omics observations (Bremen, 21-23. Feb. 2018). *Atlantos Newsletter* 1(3): 17 (2018).
- Port, J.A. et al. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology* 25(2): 527-541 (2016).
- Rabinovici, S.J. et al. Economic and health risk trade-offs of swim closures at a Lake Michigan beach, *Environmental Science and Technology* 38(10): 2737-2745 (2004).
- Rees, H.C. et al. Review: The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51(5): 1450-1459 (2014).
- Schmelzle, M.C. and A.P. Kinziger. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. *Molecular Ecology Resources* 16(4): 895-908 (2016).
- Scharff, R.L. Economic burden from health losses due to foodborne illness in the United States, *Journal of Food Protection* 75(1), 123-131 (2012).
- Southwick Associates. Comparing NOAA's recreational and commercial fishing economic data. [https://asafishing.org/uploads/Comparing\\_Recreational\\_and\\_Commercial\\_Marine\\_Fishing\\_Data\\_Report\\_May\\_2013.pdf](https://asafishing.org/uploads/Comparing_Recreational_and_Commercial_Marine_Fishing_Data_Report_May_2013.pdf) (2013).
- Stephens Z.D., Lee S.Y., Faghri F., Campbell R.H., Zhai C., Efron M.J., Iyer R., Schatz M.C., Sinha S, Robinson GE. Big data: Astronomical or genomics? *PLOS Biology* 13(7): e1002195 (2015).
- Stepien, C.A, Snyder M.R., Elz, A.E. Invasion genetics of the silver carp *Hypophthalmichthys molitrix* across North America: Differentiation of fronts, introgression, and eDNA metabarcoding detection. *PLOS One* 14(3): e0203012 (2019).

- Stulberg, E., Fravel, D., Proctor L.M., Murray D.M., LoTempio J., Chrisey L., Garland J., Goodwin K., Graber J., Harris M.C., Jackson S., Mishkind M., Porterfield D.M., Records. An assessment of US microbiome research. *Nature Microbiology* 1:15015 (2016).
- Thomsen, P.F. et al. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLOS One* 7(8): e41732 (2012).
- Thomsen, P.F. and E. Willerslev. Environmental DNA - an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* 183: 4-18 (2015).
- Thomsen, P.F. et al. Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. *PLOS One* 11(11): e0165252 (2016).
- Thompson, L. et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551: 457-463 (2017).
- UNIG. Ray, J. Strom, M., Goodwin K.D. (eds). Implementation of Environmental DNA (eDNA) as a Tool for Ecosystem-Based Fisheries Management. in press. (2020).
- Watts, A.W., Miksis-Olds, J. The Oceans as a Living Resource: Environmental DNA and Acoustics for Detecting Marine Life. Discussion Paper for the National Conference on Marine Environmental DNA. [https://phe.rockefeller.edu/eDNAMarine2018/papers/Watts-Miksis-Olds\\_paper\\_Ocean\\_as\\_a\\_Living\\_Sensor\\_112618.pdf](https://phe.rockefeller.edu/eDNAMarine2018/papers/Watts-Miksis-Olds_paper_Ocean_as_a_Living_Sensor_112618.pdf) (2018).
- Werner, C., Armor, J., Matlock, G. 3 April 2018. Emerging Technologies for NOAA Ocean Research, Operations, and Management in an Ecosystem Context. Response from NOAA to a Report from the NOAA Science Advisory Board. Preliminary Report. [ftp://ftp.oar.noaa.gov/SAB/sab/Meetings/2018/April/EmergingTechnologies\\_Preliminary\\_Response\\_SAB\\_3April2018.pdf](ftp://ftp.oar.noaa.gov/SAB/sab/Meetings/2018/April/EmergingTechnologies_Preliminary_Response_SAB_3April2018.pdf) (2018).
- Yamamoto, S. et al. Environmental DNA as a 'snapshot' of fish distribution: a case study of Japanese jack mackerel in Maizuru Bay, Sea of Japan. *PLOS One* 11(3): e0149786 (2016).
- Yoder, J., Fluharty, D., Leong, J-A. Emerging Technologies for NOAA Ocean Research, Operations, and Management in an Ecosystem Context. Draft Report of the Ecosystem Science and Management Working Group (ESMGWG) to NOAA Science Advisory Board. [https://sab.noaa.gov/sites/SAB/Meetings/2016%20Documents/November%202016%20Documents/Emerging%20Technologies%20Report\\_final%20with%20title%20page.pdf](https://sab.noaa.gov/sites/SAB/Meetings/2016%20Documents/November%202016%20Documents/Emerging%20Technologies%20Report_final%20with%20title%20page.pdf) (2016).