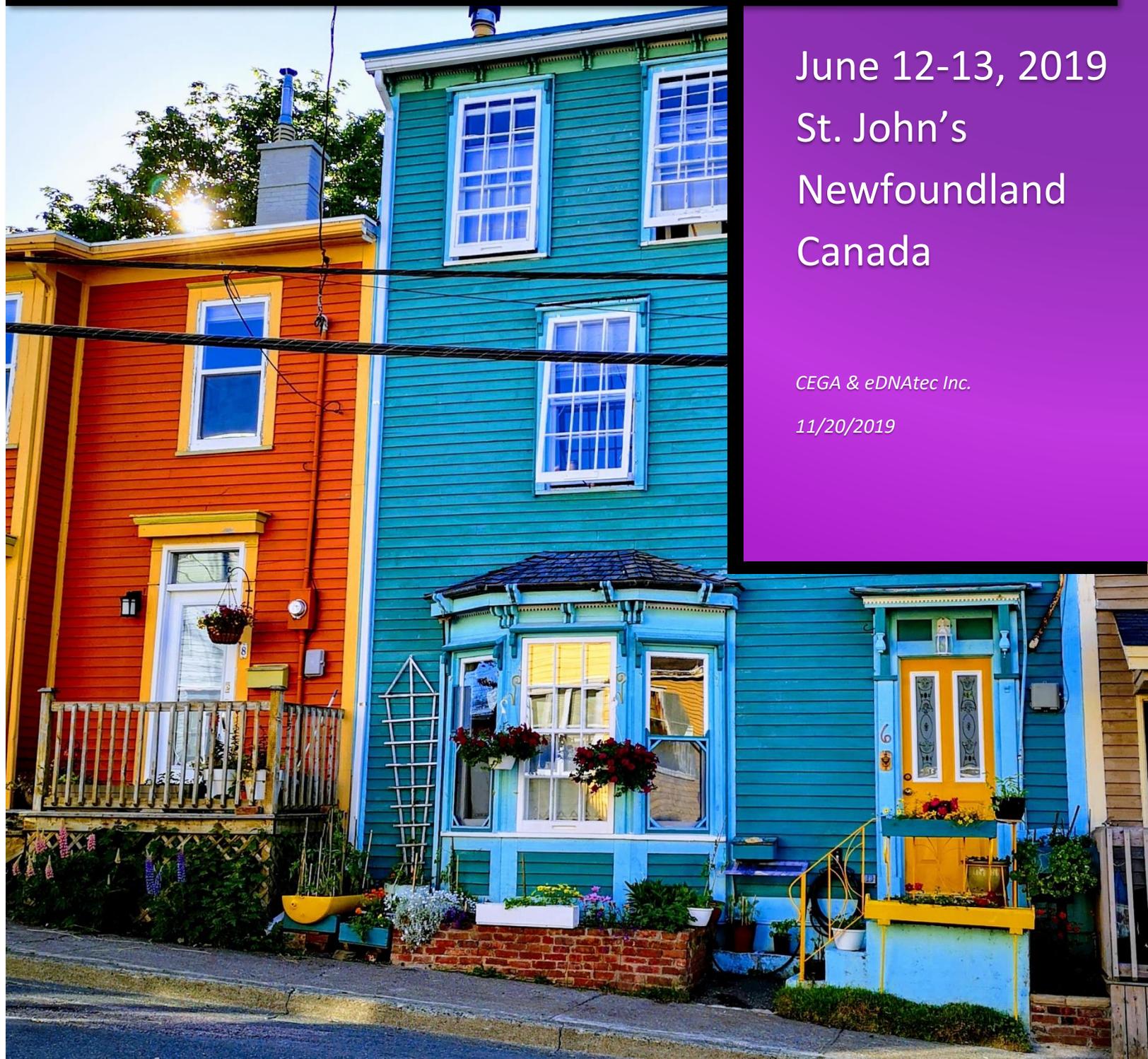


2019 International Workshop on Environmental Genomics *Scalable Solutions*

June 12-13, 2019
St. John's
Newfoundland
Canada

CEGA & eDNAtec Inc.

11/20/2019



IWEG 2019 SPONSORS



GenomeAtlantic

illumina[®]



Hibernia

CONTENTS

EXECUTIVE SUMMARY 1

INTRODUCTION..... 2

SESSION SUMMARIES 4

APPENDIX 1. Survey responses from IWEg participants 10

EXECUTIVE SUMMARY

Participants from around the world joined the 4th International Workshop on Environmental Genomics (IWEG) in St. John's to discuss developments in the field of environmental genomics and the scalability of the technology. Members represented various sectors, including the oil and gas industry, environmental consulting companies, regulatory agencies, and academia. Presentations focused on existing nation-wide research programs, technological innovations, and the opportunities created by these innovations for larger scale projects. The challenges to broad-scale implementation of environmental genomics, namely DNA metabarcoding of eDNA, and potential solutions to those challenges were central to the discussion. Other themes from this year's event included the importance of robust sampling designs, standardization of genomics techniques, and advancing the adoption of these tools by regulatory agencies. Notably, seven oil and gas companies (Chevron, Shell, Total, Equinor, ExxonMobil, Hess, ENI) in the International Consortium for Environmental Genomics (ICE-G) Industry Sub-Group have advanced the Joint Industry Program (JIP) which fund projects that will contribute to the development of guidelines for the application of environmental genomics in environmental management activities in the oil and gas industry.

INTRODUCTION

The 2019 International Workshop on Environmental Genomics (IWEG), “Scalable Solutions”, was held in St. John’s, NL, Canada on June 12th and 13th. Over 50 participants from 8 countries were in attendance and represented environmental regulatory and consulting agencies, industry, and academia. The vast majority of participants (92%) had experience working with environmental genomics tools or were leading their own environmental genomics research programs (see Appendix for these and other survey results).

Continuing from past workshops and building momentum towards regulatory acceptance, IWEG 2019 focused on the increasing scale of environmental genomics studies. The workshop was organized into two main sessions: National Programs and Technology, Trends & Implications. Roundtable discussions followed each session and several key themes emerged:

- National programs adopting environmental genomics tools provide examples of early large-scale eDNA projects. These projects benefit from the existing program infrastructure for traditional monitoring methods, but they also face the challenge of adopting new methods when other methods are already established. Lessons learned from these programs provide good guidance for future large-scale programs.
- Regulatory agencies are slow to adopt eDNA but are making moves in that direction. Environmental DNA research has been expanding rapidly and regulatory agencies have taken note. Some agencies have developed working groups or panels to evaluate the potential of environmental genomics. They are taking promising steps towards integrating these techniques into their policies and establishing best practices.
- Sample collection is often the limiting step in large-scale studies. As laboratory and bioinformatic technologies advance, the time from sample collection to delivering results is decreasing rapidly. However, collecting samples in the field remains challenging and time consuming, and this will likely be the biggest hurdle as projects scale up.

- Sampling designs should be carefully planned and optimized prior to implementation in large-scale projects. This will be important to optimize costs over large spatial and temporal scales. Pilot projects are recommended to inform sampling designs for specific study goals and environments prior to beginning large projects.
- Standardization is an important goal in the environmental genomics field and essential to scalability. This has been a common theme at previous workshops, and new means of standardization were considered this year. Previously, standardizing methodologies and validating protocols across research groups was suggested; this year, agreement in results across research groups using different methodologies was suggested as a means of demonstrating reliable results.
- Environmental genomics can move beyond taxonomic lists to capitalize on other data sources and the rich information content found in DNA. Environmental monitoring can benefit from resources such as trait databases that link the rich taxonomic information obtained from eDNA metabarcoding with associated ecosystem roles through functional traits, ecological networks, and food webs. The rise of deeper sequencing technologies (e.g., the Illumina Novaseq platform) facilitates eDNA analyses that go beyond simple metabarcoding, such as querying for functional genes through metagenomics, metatranscriptomics, as well as tracking biodiversity at the population level. Leveraging genomic data beyond traditional taxonomic lists will expand the applications of environmental genomics and the utility of its deliverables.

SESSION SUMMARIES

Day 1: National Programs

The first day of the workshop began with keynote speaker Carol Stepien (Pacific Marine Environmental Laboratory, NOAA, USA). She set the theme of “scalable solutions” by first providing context on environmental genomics processes and terminology. She then reviewed ongoing DNA metabarcoding studies, including an ecosystem study of the effects of ocean acidification, an analysis of hydrothermal vent and methane seep communities, and a population genetics study of two mussel species. Her organization is taking steps to scale up their efforts by integrating eDNA into surveys in combination with other methods, such as unmanned platform technologies and traditional sampling to ground truth eDNA. Several points were mentioned in the keynote presentation that were later reiterated throughout the workshop: the need to expand industry partnerships, the importance of improving data access, and finding ways to standardize field, lab, and bioinformatic steps in the DNA metabarcoding workflow.

The following two sessions outlined some of the national programs currently using environmental genomics. Presenters highlighted successes, difficulties, and paths forward for the integration of eDNA technologies into their respective organizations.

The first session began with Donald Baird (Environment and Climate Change Canada & University of New Brunswick, Canada) who introduced the Metagenomics-Based Ecosystem Biomonitoring project. This “Ecobiomics” project adopts a team-based approach to build a standard environmental genomics platform across government agencies to better assess, monitor, and remediate microbial and invertebrate biodiversity with the aim of improving water quality and soil health. Erik Pilgrim (United States Environmental Protection Agency) showcased his work to incorporate eDNA metabarcoding into National Aquatic Resource Surveys to complement existing biomonitoring programs. Cathryn Abbott (Fisheries and Oceans Canada) demonstrated how various end-user needs can be unified in her organization using environmental genomics, and the importance of standardized and scalable technologies that

can dynamically shift based on priorities and funding. She also provided examples of time-consuming, and costly, traditional survey methods that could benefit from the implementation of eDNA-based monitoring. Jon Thomassen Hestetun (Norwegian Research Centre, NORCE) gave the initial results and next steps of METAMON (High-throughput Metabarcoding of Eukaryotic Diversity for Environmental Monitoring of Marine Sediments), a project aimed at optimizing metabarcoding protocols for environmental monitoring of marine benthic communities on the Norwegian Shelf.

Richard Lance (Environmental Laboratory, United States Army Engineer Research & Development Center) commenced session two, discussing the use of a crowd-sourced and open-access database, “eDNAtlas”, for collaborative eDNA sampling and archiving. The data are used to meet management and conservation goals across multiple departments. Shauna Baillie (Fisheries and Oceans Canada) reviewed the exponential growth of eDNA research in recent years and the paradigm shift in how we study biodiversity. This momentum led DFO to establish the National DFO eDNA Technical Working Group for better communication between scientists, and to facilitate the move towards best practices. Willie Duncan (Scottish Environmental Protection Agency) presented the challenges to eDNA technologies receiving formal adoption from management, especially in circumstances where pre-existing biomonitoring methods are already in place. Eric Stein (Southern California Coastal Water Research Project, United States) examined challenges with bioassessment indices, and how they are best adapted to account for differences between conventional biomonitoring approaches and new molecular methods. Renovating existing bioassessment indices, rebuilding new indices for molecular data, and creating “cross-walks” may be necessary for the incorporation of these new data sets and comparison with older datasets.

A recurring subject throughout the first day of presentations was the difficulties impeding adoption of eDNA into national programs that already have established biomonitoring methods. Industry adoption has been slow and there is some hesitation from regulators to embrace this technology, however funding initiatives are in place to support continued research. The need for standardization was also mentioned often during both the presentation and discussion periods, with some seeing it as a key element of scalability.

Collaboration and sharing of knowledge were also recognized as important components to the success of national programs.

Day 1: Discussion

For the first discussion session participants were asked two questions: “What do you see as the key elements of “scaling up” environmental genomics studies?” and “Hundreds of small-scale environmental genomics studies have now been published. What are some of the barriers to enacting large-scale studies?”

The key elements of “scaling up” are often specific and tailored to the goals of each project; however, several common themes emerged throughout the discussion. Longer-term studies are critical for scaling up because baseline molecular data are needed to document ecological changes, and by archiving samples, these longitudinal studies allow for future reanalysis with new technologies. Organizations can increase the scale of their environmental genomics studies by leveraging data between projects and taking advantage of collaborations. Furthermore, projects can scale up in stages by beginning with a small round of intense sampling and using the data to help inform sampling plans across larger geographical or temporal scales. Lastly, reporting and presenting success stories to industry is essential to attaining approval to scale up existing projects.

There are several barriers to administering large-scale studies but a primary concern is the need for substantial long-term financial support because funding models are typically not compatible with studies of this magnitude. This could be addressed through funding of centralized facilities dedicated to these types of large-scale environmental genomics projects that support research groups across institutions and industries. In turn, focused resources and concentration of expert knowledge at centralized facilities will also contribute to development of standards and other technological advancements. Another barrier to large-scale studies is public relations: we need to foster public interest in environmental genomics and engage people by translating complex data so that it can be shared with a wider audience.

DAY 2: Technology, Trends, & Implications

On the second day, three sessions covered trends in technology related to environmental genomics and the implications of these trends as studies scale up. One of the largest challenges to scaling up studies is the collection of samples. It is difficult to improve the efficiency of sample collection because it is often expensive, time-consuming, and can be subject to unpredictable environmental conditions. However, examples of creative solutions to overcome these challenges were presented, including the STREAM (Sequencing the Rivers for Environmental Assessment and Monitoring) project, described by Merhdad Hajibabaei (University of Guelph & CEGA, Canada), that will leverage citizen scientists to collect benthic samples across Canada for environmental genomics analysis. Hajibabaei also showcased another promising approach using new ultra-deep sequencing at CEGA (Centre for Environmental Genomics Applications) to characterize the entire biodiversity in a sample including various bioindicator groups as well as commercially important species in a single test. This approach will maximize recovery of biological information in a sample.

As studies scale up, they will also cover larger geographical areas and new environments. Environmental genomics analyses in new environments often present new challenges, especially if the area is not well studied. Thomas Dahlgren (NORCE) discussed the limitations of reference database coverage in the North Sea, considered a best-case scenario since the area is relatively well-studied. This suggests that more work is needed to build reference databases for marine environments and to achieve high levels of taxonomic identification, especially in under-studied areas. David Coté (Fisheries and Oceans Canada) discussed some preliminary results from a deep ocean environment in the Labrador Sea, an example of an under-studied marine ecosystem. He highlighted the challenges of sampling in the deep ocean and interpreting the results with limited taxonomic information.

New sequencing technologies allow studies to scale up by sequencing at greater depths and by decreasing the turn-around time from sampling to results. Kahlil Lawless (Illumina Inc., Canada) gave an overview of Illumina's current sequencing platforms and recent developments in their capabilities, including increased sequencing depth alongside longer read lengths

(NovaSeq) and increased portability of sequencing devices (iSeq). Mehrdad Hajibabaei (University of Guelph & CEGA, Canada) demonstrated the power of deep sequencing by showing that metabarcoding on the NovaSeq improved target species detection probabilities compared to shallow targeted sequencing with species-specific primers or species-specific qPCR. Tiffany Simpson (Curtin University, Australia) showed the results of using the iSeq instrument on a vessel at sea, decreasing the turnaround time for sample processing and generating sequencing results while in the field. Additionally, the MinION platform from Oxford Nanopore has been used for microbiome sequencing projects at the point of need, as demonstrated by Richard Tennant (University of Exeter, United Kingdom) and Thomas Merzi (Total S.A., France). Both the Illumina iSeq platform and the MinION demonstrate the portability of new sequencing instruments and their ability to be used in the field to generate sequencing data quickly. Turnaround time can also be improved with faster bioinformatics, where most steps can easily be parallelized for faster computational times, as demonstrated by Greg Singer (CEGA, Canada).

As technology matures, environmental genomics methods are now being used by more people and a broader range of stakeholders. This session showcased projects conducted by both regulators and industry members and included routine monitoring, exploration, and impact assessments in the marine environment. For example, Nicholas Jeffery (Department of Fisheries and Oceans Canada) discussed how the Canadian government is incorporating eDNA technologies as a non-invasive sampling tool into their Marine Protected Areas monitoring. In the oil and gas industry, Nicolas Tsesmetzis (Shell International Exploration and Production Inc., United States) described how environmental genomics is being used for exploration by connecting microbial community composition and metabolic processes to reveal the type of hydrocarbons found in the area. Marc Skinner (Stantec Inc., Canada) presented an overview of experiments aimed at developing eDNA sampling techniques to monitor fish species at tidal energy development sites. As environmental genomics projects grow in scale and the userbase broadens, technologies are developing to meet emergent challenges. Environmental genomics continues to expand with many new applications and opportunities for genetic information to both complement and enhance conventional ecological surveys.

DAY 2: Discussion

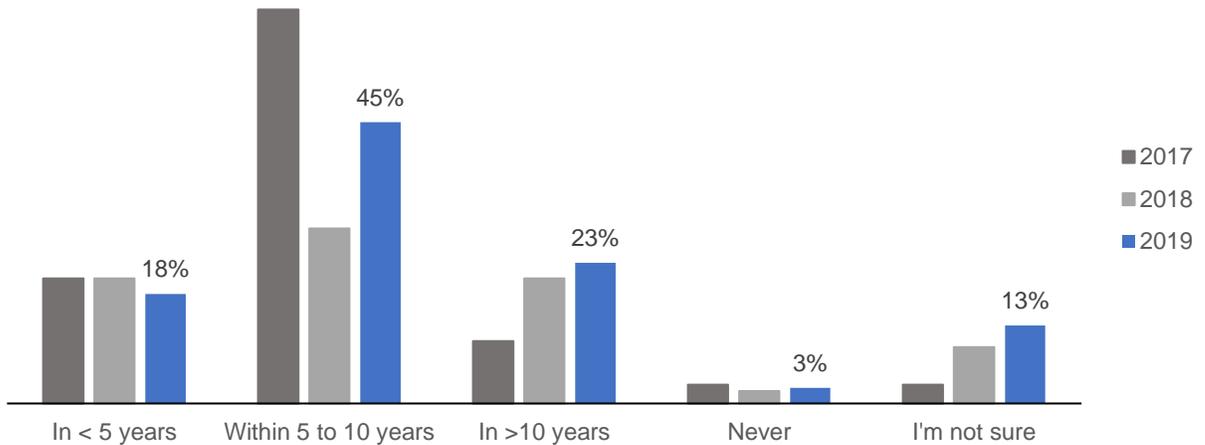
During the discussion period on the second day, the questions participants addressed were “What are the more challenging environments to conduct large-scale eDNA analysis and how do monitoring needs differ between these environments?” and “What are the key outcomes you’d like to see from large-scale studies?”.

Locations in terrestrial, marine, and freshwater environments all present challenges in sampling, suggesting that challenges extend across ecosystems and habitats. Some examples of difficult sampling areas across ecosystems include polar regions, areas far from infrastructure, areas in conflict zones, environments with low DNA signal strength (e.g. deep ocean), environments with a lot of DNA inhibitors (e.g. tannic waters), and environments that are highly variable or unpredictable (e.g. the hyporheic zone of streams). In most of these cases, obtaining the samples is the most difficult step. However, in some environments, the lack of reference database coverage is also a major challenge.

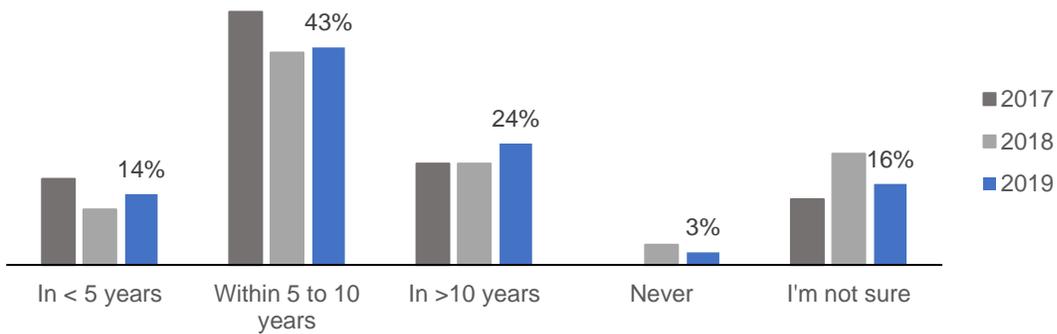
There are several key outcomes that practitioners are looking for as studies scale up. With unlimited potential goals for large-scale environmental genomic surveys, studies must continue to have focused designs that answer the specific research or monitoring questions efficiently and minimize the amount of sampling and analysis required. Field pilot projects are a useful first step in large-scale studies to inform these optimal sampling designs for detecting target species or trends of interest. Large-scale studies should provide a better understanding of natural spatial and temporal variation in communities to increase our power to detect aberrations from natural and anthropogenic disturbances. Finally, studies should move from taxonomic lists as the final result of environmental genomics, towards approaches that capitalize on existing databases and the rich information content in DNA that can be uncovered by deeper sequencing technologies (e.g., Novaseq). Some of these approaches included exploring phylogenetic and functional relationships, metatranscriptomics, population genetics, and ecological networks or food webs constructed by supplementing DNA taxonomic lists with information from occupancy modeling or existing traits databases.

APPENDIX 1. Survey responses from IWEG participants

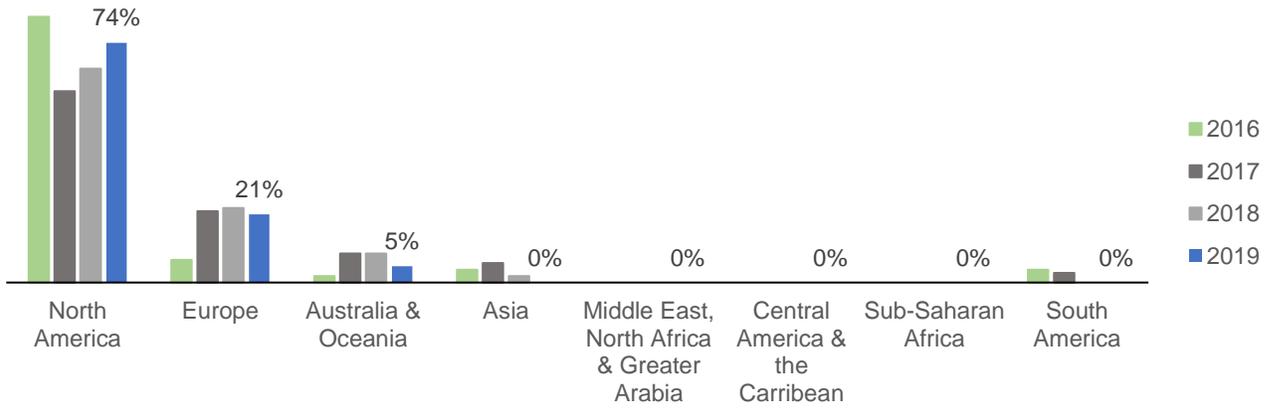
1) When do you see environmental genomics becoming the primary approach to characterize and monitor biodiversity in the *marine/aquatic* environment?



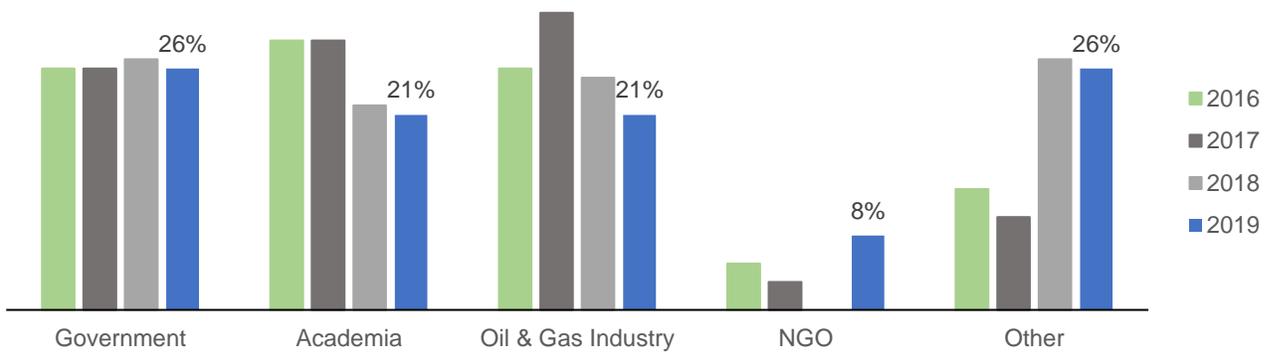
2) When do you see environmental genomics becoming the primary approach to characterize and monitor biodiversity in the *terrestrial* environment?



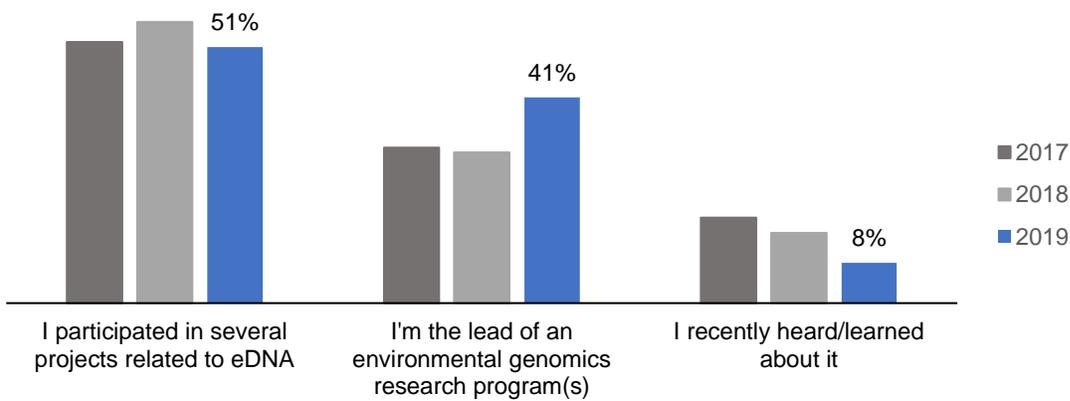
3) Which region are you from?



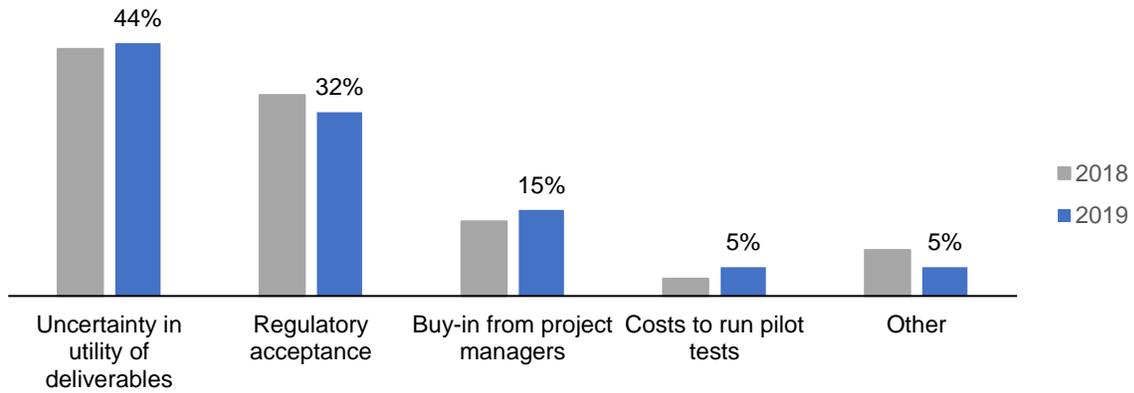
4) What group do you represent?



5) How familiar are you with "Environmental Genomics"?



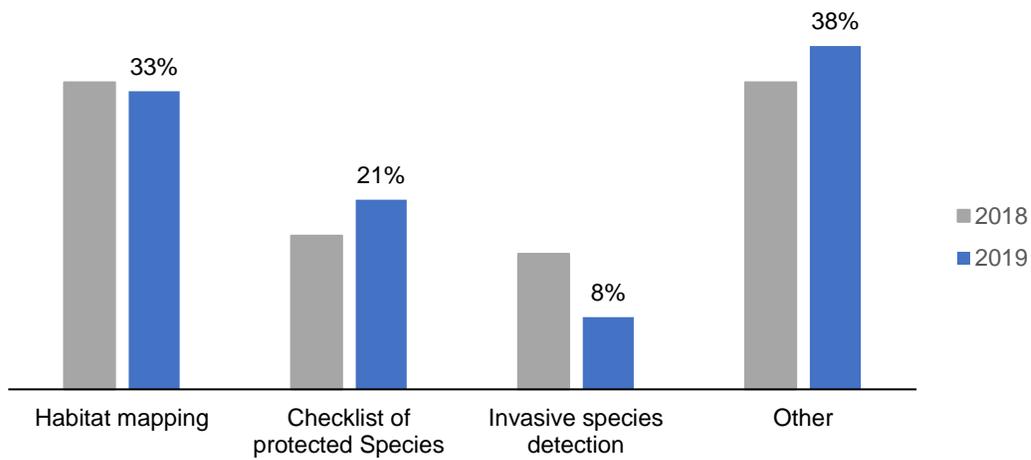
6) What are the biggest barriers to introducing environmental genomics in new and existing environmental assessment and monitoring efforts?



Comments:

- *Inability to track abundance and life stage.*

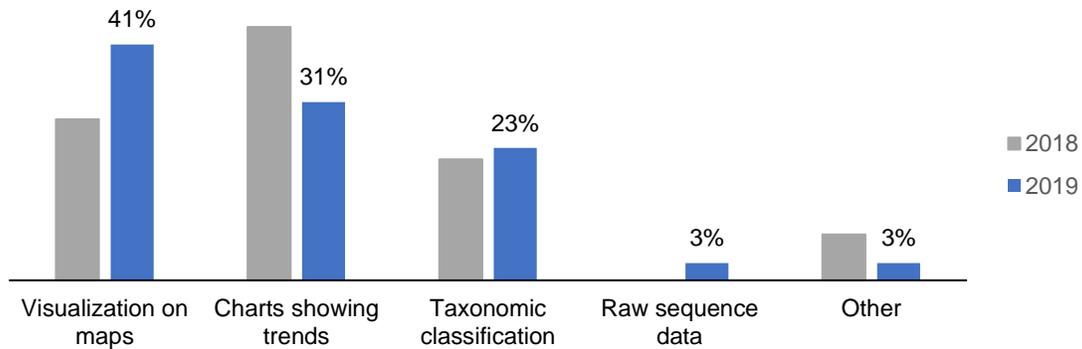
7) From a regulatory perspective, what is the highest priority?



Comments:

- *Ecosystem health*
- *Impact monitoring*
- *Seasonal trends*
- *Comparability with historic data*

8) What type of reporting do you find most useful?



Comments:

- *Incorporation in Environmental Assessment outputs*

9) What do you hope to get out of this workshop?

Comments:

- *Update on state of the art and trends for eDNA and research gaps*
- *Opportunities for networking and for collaboration to advance the science of environmental genomics*
- *To learn directions and capabilities of emerging technology*
- *Productive dialogue about where we should focus our efforts to take the next steps towards regulatory approval and industry adoption*
- *Awareness of how to move towards implementation of new technologies on larger scale*