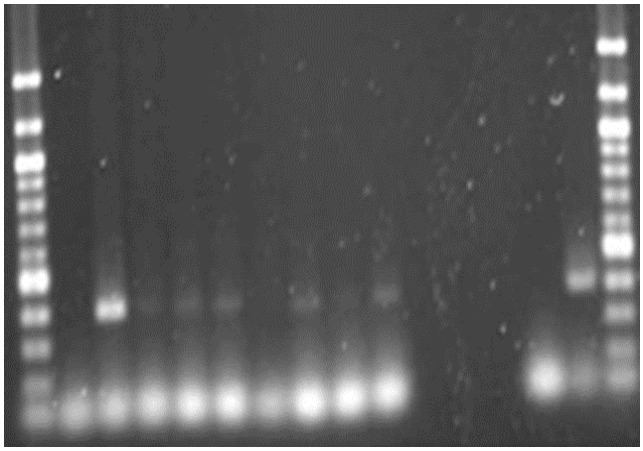
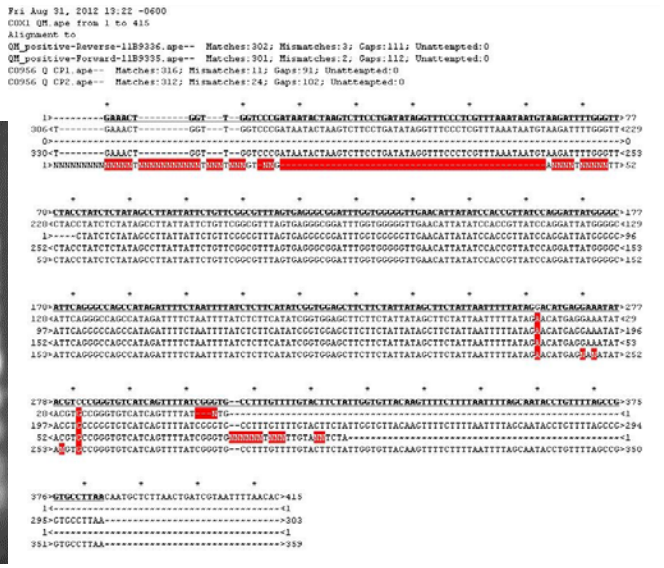
RECLAMATION

Managing Water in the West

Microbiome Analysis of Lake Mead

Research and Development Office
Science and Technology Program
Final ST-2016-5385-1





U.S. Department of the Interior
Bureau of Reclamation
Research and Development Office

September 2016

Mission Statements

The U.S. Department of the Interior protects America's natural resources and heritage, honors our cultures and tribal communities, and supplies the energy to power our future.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188
T1. REPORT DATE Sept 2016	**T2. REPORT TYPE** Research	**T3. DATES COVERED** 2016 FY
T4. TITLE AND SUBTITLE Microbiome Analysis of Lake Mead Using Microbiome Analysis to Assess the Impact of the Dreissenid Mussel Populations on the Ecology of Lake Mead		**5a. CONTRACT NUMBER** RY1541ZQ201615385 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER** 1541 (S&T)
6. AUTHOR(S) Jacque Keele, jkeele@usbr.gov, 720-930-1056 Denise Hosler, dhosler@usbr.gov, 303-250-9166		**5d. PROJECT NUMBER** 5385 **5e. TASK NUMBER** **5f. WORK UNIT NUMBER** 86-68560
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Bureau of Reclamation, Technical Service Center, Hydraulic Investigations & Lab Services		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Research and Development Office U.S. Department of the Interior, Bureau of Reclamation, PO Box 25007, Denver CO 80225-0007		**10. SPONSOR/MONITOR'S ACRONYM(S)** **R&D: Research and Development Office** **BOR/USBR: Bureau of Reclamation** **DOI: Department of the Interior** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** 2016-5385-1
12. DISTRIBUTION / AVAILABILITY STATEMENT **Final report can be downloaded from Reclamation's website: https://www.usbr.gov/research/**		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT *(Maximum 200 words)* Next generation DNA sequencing (NGS) is allowing researchers to study and understand the ecology of the microbial communities from a wide range of locations and conditions. For this scooping project samples were collected from Lake Mead, NV and the Salton Sea, CA for microbial analysis. The DNA was extracted and sent for NGS analysis. The goal of this project was to plan and design a NGS study. Lake Mead was selected as one of the sites because of its large invasive Dreissenid mussel population. The Salton Sea was selected because of its unique chemistry and environment. Once the NGS analysis is complete a list of the microbial organisms present at each site will be compiled. The data collected in this scooping project will be used in future projects to understand how the microbiome is impacted by both the presence of invasive mussels and environmental changes.

15. SUBJECT TERMS Next generation sequencing, microbiome, DNA

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Jacque Keele
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	U	20	19b. TELEPHONE NUMBER 720-930-1056

PEER REVIEW DOCUMENTATION

Project and Document Information

Project Name Microbiome Analysis of Lake Mead WOID Z5385

Document Microbiome Analysis of Lake Mead

Document Author(s) Jacque Keele Document date September 2016

Peer Reviewer Yale Passamaneck

Review Certification

Peer Reviewer: I have reviewed the assigned items/sections(s) noted for the above document and believe them to be in accordance with the project requirements, standards of the profession, and Reclamation policy.

Reviewer _____ Date reviewed 9/16/2016
(Signature)

Disclaimer

This project not endorse any particular company for the analysis of NGS samples.

Acknowledgements

The authors of this research would like to thank the Research and Development Office for supporting and providing funding for this study. The authors wish to thank all current members of the RDLES laboratory for their support and assistance. In particular, special thanks to Yale Passamaneck for collecting samples at Lake Mead and the Salton Sea. We also wish to thank the reviewer of this document Yale Passamaneck.

Notices

None

Acronyms and Abbreviations

eDNA- Environmental DNA
NGS- Next generation DNA sequencing
RDLES-Reclamation Detection Laboratory for Exotic Species

Executive Summary

Over the last few years, the ability of researchers to perform analysis of the microbiome (the microbial community) by next generation sequencing (NGS) has given a fuller understanding of microbial communities from a wide range of environments. With NGS, researchers are able to inventory all of the organisms present in an environment (viruses, bacteria, protozoa, etc.). This technology is important because many of these organisms are difficult to culture in the laboratory, and NGS can be a way of determining what organisms are present in the sample. The use of microbiome analysis to assess the impact of an invasive organism, pollutant, or environmental change needs to be explored so that researchers at Reclamation can come to understand, appreciate, and take advantage of this technology.

Two sites were selected for this initial NGS project: Lake Mead and the Salton Sea. For the last 10 years Lake Mead has been at the epicenter of the *Dreissena bugensis* (quagga) mussel infestation in the western United States. The Salton Sea is an environment where there is high salinity and where the water level is continuing to decease. For both sites, underlying the ecology of these bodies of water are the micro-organisms that inhabit the system. The microbial community is important because of the roles it plays in decomposition, recycling, and as a food source for higher invertebrates and vertebrates. Understanding the microbiome can also help improve the understanding of the chemical recycling pathways (carbon and nitrogen for example) and biogeochemical processes that occur in bodies of water.

This project had several goals. First, to plan and design a microbiome study. Second, collect samples and extract DNA for the analysis. Finally, to analyze the DNA samples by NGS sequencing to create a list of organisms that are present in the water. The first two goals (sample collection and DNA extraction) were accomplished in 2016. The third goal is ongoing because the sequencing company has not yet completed the NGS sequencing and data analysis. Once those results are available they will be added as data files for this report.

From each of these steps lessons were learned that will be applied to future NGS studies. First, the design and planning of a NGS project can be either very simple or complex. For this initial project it was decided to keep the design simple by analyzing samples from only two sites. Future projects could involve multiple sites and time points to assess the microbiome. Second, sample collection and preservation is key. The way that the samples are collected will impact the outcome because the size of the filter used will select for different sizes of organisms. Also, if the samples are not preserved the DNA will be degraded before the DNA extraction is even performed. These lessons that were learned will help in carrying out future NGS project.

The next step is to use the knowledge and experience gained by this project to carry out future NGS projects. For fiscal year 2017, the author submitted a research proposal for to perform NGS on the waterbodies of the Salt River Project in Arizona. This system offers a unique opportunity for research because there is an emerging quagga mussel population in two of the reservoirs of the system. Other future projects will also make use of the NGS technology to address many different research questions.

Contents

Tables

Main Report

Introduction

The use of next generation DNA sequencing (NGS) technology to assess and catalog microbiome populations has been done on a wide range of samples from environmental to human gut analysis [1]. A great deal of research has been performed to study the role that the microbiome plays in human health [2]–[9]. There have also been many studies that analyze the microbiomes of environmental samples [10]–[19]. The microbiome can encompass all of the bacteria, viruses, fungi, and archaea that live in a system. Over the last few years the number of studies that use NGS has increased. With the ability to catalog the organisms that are present in a sample, the next step for this technology will be to use these data to create models and further our understanding of the interplay of the microbiome with the environment [20].

In an effort to start to use this technology, two sites were selected: Lake Mead, NV and the Salton Sea, CA for NGS analysis. From each of these sites, DNA samples were collected and extracted. The DNA was sent to a commercial company for NGS analysis. Once the NGS company completes the analysis the data will be added to this report.

These sites were selected for several reasons. Lake Mead is the epicenter for the invasive *Dreissena bugensis* (quagga mussel) in the Western United States. The presence of these mussels has altered the ecology of Lake Mead. The Salton Sea is a unique environment with high salinity. The impact of invasive mussels at Lake Mead and climate change at the Salton Sea for these two very different bodies of water is of interest to researchers at Reclamations Detection Laboratory for Exotic Species (RDLES). By collecting NGS samples from these two sites in 2016, it is hoped that future NGS experiments in 2017 and beyond will be able to build on both the lessons learned in the process and the results that are gathered during this scooping project.

There are many studies where NGS has been used to analyze the microbiome. To give an idea of how many NGS studies of microbiomes have been performed a key word search of NCBI PubMed revealed over 24,000 publications. Many of these studies are on human and animal microbiomes of the gut and stomach. There is an ongoing human microbiome project that had the goal of identifying and characterizing the microbial flora of health and sick individuals. One of the major attractions of NGS analysis of microbial samples is that many microbes cannot be easily or cannot be grown in the laboratory setting. With the advent of NGS it was possible to catalog organisms that could not be grown in the laboratory. Microbiome research has opened a window into a very complex world of bacteria, viruses, and other prokaryotic organisms.

There are several research questions that can be answered by conducting a microbiome analysis of Lake Mead and other bodies of water in the western United States. For example, how diverse is the current microbiome at Lake Mead? Is there a wide range of organisms or has it narrowed to a few classes due to the quagga mussel infestation? The NGS analysis will enable researchers to have a fuller understanding of the impact of the quagga mussel on Lake Mead. This technology will allow researchers to compare the microbiome of many different sites to determine if there are organisms that are absent or over represented in waters where quagga mussels are present. This preliminary project was performed to allow RDLES researchers to learn the best way to conduct a NGS project.

Experimental Design

There are three phases to this study. First, plan and design the microbiome study. Second, carry out sample collection, DNA extraction, and next generation sequencing. Finally, the analysis of the NGS data will be used to build a picture of microbiomes in western waters. This will be a descriptive analysis of the microbiome of Lake Mead and the Salton Sea. Both of these locations have been heavily sampled and monitored for multiple years, so there is a wealth of environmental and chemical data that can be related to the microbiome results. Hopefully it will be possible to relate the microbiome results to the environmental data that has previously been collected. This aspect of the data analysis is beyond the scope of this project and will be an ongoing part of future NGS projects.

The first task in this project was to plan and design the microbiome study. One of the most important tasks was to identify a company that could perform the NGS sequencing analysis. This involved online searches and contacting several different companies (both commercial and university facilities) to obtain quotes and determine which would best serve the needs of this project. Eventually, a consulting company (Genohub) that could act as a guide was identified. By working with this company it was possible to identify a commercial company who could perform the NGS testing and also analysis. Based on recommendations from the consulting company it was decided to analyze the 16S V4 ribosomal region to obtain sequences to identify the different microbes in the sample. The read lengths would be 2 x 300 (paired end). Both the forward and reverse sequence will be analyzed. They also guaranteed number of pass filter paired-end reads per sample as 100,000. Thus, this initial project NGS project will create a large amount of data.

It was also decided to have the sequencing company perform the initial data analysis. This is the most important step of the NGS process and understanding the best bioinformatics programs and methods to use to analyze the data is key to having usable results. The commercial company performed basic de-noising and chimera check of the data. They will also perform a basic taxonomic analysis of the data.

Sample Collection
At Lake Mead and the Salton Sea, a single sample was collected by the boat ramp. Samples were collected by filtration using a 0.45 µm GN-6 Grid cellulose filter (Pall 66068). The filter was then placed in RNA*later*® (ThermoFisher Scientific, AM7030) for DNA stabilization and storage. The samples were then shipped back to the RDLES Laboratory for further processing. Upon arrival at RDLES the samples were refrigerated until the DNA extraction could be performed.

DNA Extraction
Measures were taken to ensure that there was no cross contamination with microbes in the laboratory and the samples. This is an important issue when performing NGS analysis [21]. Prior to extracting the DNA the forceps that were used to handle the filters were sterilized in bleach. At all stages in the process nitrile gloves were worn and precautions taken to avoid cross contamination. The quality control/quality assurance steps that RDLES takes for all samples were followed.

The filters were removed from the RNA*later*® with forceps, and then cut into pieces. These pieces were placed in a 2.0 mL Eppendorf tube and the Qiagen DNeasy Blood and Tissue Kit (69504) was used to extract the DNA. The samples were incubated overnight in the ATL buffer and proteinase K reagents. After the extraction was completed the $OD_{260/280}$ was taken using the spectrophotometer for both samples to determine the DNA concentration (Table 1).

Table 1: Sample OD readings and concentrations

Sample	OD 260	OD 280	Ratio 260/280	Concentration (ng/µL)
Lake Mead	0.0138	0.0101	1.3	0.69
Salton Sea	0.0678	0.0484	1.4	3.39

DNA concentration was determined using the formula:
OD_{260} X conversion factor = µg/mL of nucleic acid
Conversion factor used was 1 OD_{260} = 50 µg/mL of dsDNA (double stranded DNA)

The purity of the DNA was determined by taking the $OD_{260/280}$ ratio. The ratios that were obtained showed that the DNA has some contamination from proteins. Because the ratios were above 1.

Once the samples were prepared they were shipped overnight to the commercial NGS company for analysis. The company has to create a library from the DNA sample and then analyze the sequences on the NGS instrument. Once they have completed this process they will provide RLDES researchers with the results.

Data Analysis
Once the lists of organisms are returned to RDLES researchers they will be analyzed. The major families of microbes present at each site will be determined. This data will be placed on the T drive.

Lessons Learned
There were several lessons learned through this project. First, the number of samples decreased over time. This was due to the cost of the analysis. For this project cost of the library creation and then the sequencing for two samples worked out to be around $440. Future projects with more samples will have larger sequencing budgets will have to go through the procurement process.

A second issue is that there are many companies that offer NGS services and they are all very similar in what they provide and in their services. However, not every company is the same in their customer service and picking the cheapest company might not give us the best results. When the consulting company (Genohub) was found it turned out to be a very helpful to the whole process. Their guidance and assistance it selecting a commercial company that could provide NGS services was very valuable. Future NGS project will most likely involve using this consulting company again.

A final lesson learned is to keep the research questions simple. At the beginning of this project it was proposed to collect many more samples from several locations and different sampling depths. Overtime, it was realized that this was not feasible for this scooping project and it was decided to simplify the experiment to two samples from two very different sites and compare the NGS results.

Future Directions
Next generation sequencing has its place in Reclamation projects. With this one method it is possible to determine all of the organisms that are present in a single sample. The power of this analysis for a wide range of projects from environmental water and soil sampling, gut content analysis, and also determination of pathogens in a water reuse sample will only continue to grow. In the coming years there will be a movement from just building catalogs of the organisms present in the sample to gaining a fuller understanding of how these microbes are interacting with each other and the environment. The use of this technology will continue to increase. Already RDLES researchers have plans for additional NGS projects to analyze samples from a wider range of water bodies.

References

[1] J. Shendure and H. Ji, “Next-generation DNA sequencing,” *Nat Biotechnol*, vol. 26, pp. 1135–1145, 2008.

[2] J. U. Scher and S. B. Abramson, “The microbiome and rheumatoid arthritis.,” *Nat. Rev. Rheumatol.*, vol. 7, no. 10, pp. 569–78, 2011.

[3] I. Nasidze, J. Li, D. Quinque, K. Tang, and M. Stoneking, “Global diversity in the human salivary microbiome,” *Genome Res.*, vol. 19, no. 4, pp. 636–643, 2009.

[4] C. J. Stewart, E. C. L. Marrs, A. Nelson, C. Lanyon, J. D. Perry, N. D. Embleton, S. P. Cummings, and J. E. Berrington, “Development of the Preterm Gut Microbiome in Twins at Risk of Necrotising Enterocolitis and Sepsis,” *PLoS One*, vol. 8, no. 8, 2013.

[5] E. a Grice and J. a Segre, “The skin microbiome.,” *Nat. Rev. Microbiol.*, vol. 9, no. 4, pp. 244–53, 2011.

[6] M. I. Smith, W. Turpin, A. D. Tyler, M. S. Silverberg, and K. Croitoru, “Microbiome analysis - from technical advances to biological relevance.,” *F1000Prime Rep.*, vol. 6, no. July, p. 51, 2014.

[7] L. K. Ursell, J. L. Metcalf, L. W. Parfrey, and R. Knight, “Defining the human microbiome,” *Nutr. Rev.*, vol. 70, no. SUPPL. 1, 2012.

[8] T. Human and M. Project, “Structure, function and diversity of the healthy human microbiome.,” *Nature*, vol. 486, no. 7402, pp. 207–14, 2012.

[9] S. R. Gill, M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. a Relman, C. M. Fraser-Liggett, and K. E. Nelson, “Metagenomic analysis of the human distal gut microbiome,” *Science (80-.).*, vol. 312, no. 5778, pp. 1355–1359, 2006.

[10] E. F. DeLong, “Microbial community genomics in the ocean.,” *Nat. Rev. Microbiol.*, vol. 3, no. 6, pp. 459–69, 2005.

[11] A. Shade, H. Peter, S. D. Allison, D. L. Baho, M. Berga, H. B??rgmann, D. H. Huber, S. Langenheder, J. T. Lennon, J. B. H. Martiny, K. L. Matulich, T. M. Schmidt, and J. Handelsman, “Fundamentals of microbial community resistance and resilience,” *Frontiers in Microbiology*, vol. 3, no. DEC. 2012.

[12] A. Barberán, A. Fernández-Guerra, B. J. M. Bohannan, and E. O. Casamayor, “Exploration of community traits as ecological markers in microbial metagenomes,” *Mol. Ecol.*, vol. 21, no. 8, pp. 1909–1917, 2012.

[13] A. Konopka, “What is microbial community ecology?,” *ISME J.*, vol. 3, no. 11, pp. 1223–30, 2009.

[14] J. C. Tang, T. Kanamori, Y. Inoue, T. Yasuta, S. Yoshida, and A. Katayama, “Changes in the microbial community structure during thermophilic composting of manure as detected by the quinone profile method,” *Process Biochem.*, vol. 39, no. 12, pp. 1999–2006, 2004.

[15] S. D. Frey, R. Drijber, H. Smith, and J. Melillo, “Microbial biomass, functional capacity, and community structure after 12 years of soil warming,” *Soil Biol. Biochem.*, vol. 40, no. 11, pp. 2904–2907, 2008.

[16] J. L. Garland, “Analysis and interpretation of community-level

physiological profiles in microbial ecology," *FEMS Microbiology Ecology*, vol. 24, no. 4. pp. 289–300, 1997.
[17] C. Sheehan, L. Kirwan, J. Connolly, and T. Bolger, "The effects of earthworm functional diversity on microbial biomass and the microbial community level physiological profile of soils," *Eur. J. Soil Biol.*, vol. 44, no. 1, pp. 65–70, 2008.
[18] T. C. Jeffries, N. J. Curlevski, M. V. Brown, D. P. Harrison, M. A. Doblin, K. Petrou, P. J. Ralph, and J. R. Seymour, "Partitioning of fungal assemblages across different marine habitats," *Environ. Microbiol. Rep.*, vol. 8, no. 2, pp. 235–238, 2016.
[19] G. Steinert, M. W. Taylor, P. Deines, R. L. Simister, N. J. de Voogd, M. Hoggard, and P. J. Schupp, "In four shallow and mesophotic tropical reef sponges from Guam the microbial community largely depends on host identity," *PeerJ*, vol. 4, p. e1936, 2016.
[20] M. K. Waldor, G. Tyson, E. Borenstein, H. Ochman, A. Moeller, B. B. Finlay, H. H. Kong, J. I. Gordon, K. E. Nelson, K. Dabbagh, and H. Smith, "Where Next for Microbiome Research?," *PLoS Biol.*, vol. 13, no. 1, 2015.
[21] S. J. Salter, M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. J. Loman, and A. W. Walker, "Reagent and laboratory contamination can critically impact sequence-based microbiome analyses," *BMC Biol.*, vol. 12, no. 1, p. 87, 2014.

Data Sets that support the final report

If there are any data sets with your research, please note:

- Share Drive folder name and path where data are stored: T Drive, ENGRLAB, HYDLAB, RDLES, Mussel Sample, 2016, Microbiome Data

- Point of Contact name, email and phone: Jacque Keele, jkeele@usbr.gov, (720) 930-1056

- Short description of the data: Data files of the sequencing results for Lake Mead and the Salton Sea
- Keywords: Next Generation Sequencing, microbiome

- Approximate total size of all files: Data pending

-

Date files: NGS results for Lake Mead and Salton Sea (to be added once the results are obtained from the NGS sequencing company).