12-2017

Surveying the Circular Rep Encoding Single Stranded (CRESS) DNA viral consortium found in invasive quagga mussels (Dreissena rostriformis bugensis) and sediments of the central Lake Michigan benthos

Nicholas Gezon
Grand Valley State University

Surveying the Circular Rep Encoding Single Stranded (CRESS) DNA viral consortium found in invasive quagga mussels (*Dreissena rostriformis bugensis*) and sediments of the central Lake Michigan benthos

NICHOLAS RYAN GEZON

A Thesis Submitted to the Graduate Faculty of Grand Valley State University

In

Partial Fulfillment of the Requirements

For the Degree of

Master of Biology

Department of Biology

December 2017

Preface:

"If you will cling to Nature, to the simple in Nature, **to the little things that hardly anyone sees**, and that can so unexpectedly become big and beyond measuring."

Rainer Maria Rilke

From Letters to a Young Poet

Acknowledgements:

I would like to thank my adviser, Dr. Kevin B. Strychar, for his huge help in designing, executing, and completing this project. Dr. Ian Hewson, Elliot Jackson, and Kalia Bistolas for helping me and hosting me during my time at Cornell University in Ithaca, NY. My committee members, Dr. Charlyn Partridge and Dr. Jim McNair, for all their help with various questions on the project. Mike Russell and my cousin Ben Gezon from White Lake Dock and Dredge for assistance with the nearshore sediment coring operation. Dr. Mark Luttenton, Justin Wegner, and the crew of the W.G. Jackson for assistance with the offshore sediment coring and quagga mussel collection. My good friend Taylor Pratt for taking me out on his boat to collect quagga mussels. Greg Chorak and Jim Den Uyl for assistance in free diving in the cold Lake Michigan water for zebra and quagga mussels in fall 2016. Our 2016 REU student, Marielys Santana, for her assistance in the sediment coring and quagga mussel sampling. Our 2017 REU student, Rebekah Bryant, for her huge help in the lab with LMQMvls1241 qPCR. Tony Fiore for help with various construction projects including the offshore gravity corer. Dr. Doug Haywick, Dr. Justin Sanders, and Paul White for their help with sediment core grain size analysis and Cs^{137} dating at the University of South Alabama. This work would not have been possible without funding support from the Michigan NASA Space Grant Consortium, Grand Valley State University Presidential Grant, Annis Water Resources Institute, The Hewson Lab, and NSF REU. Finally, I need to thank my parents for their endless love and support through this process.

Table of Contents:

List of Tables:

List of Figures:

Abbreviations:

AWRI	Annis Water Resources Institute
bp	Base pairs
CRESS	Circular Rep Encoding Single Stranded
C_t	Concentration threshold
DGC	Deepwater Gravity Corer
LM	Lake Michigan
ML	Muskegon Lake
nt	Nucleotides
PCR	Polymerase Chain Reaction
QM	Quagga mussel
qPCR	Quantitative Polymerase Chain Reaction
RT- qPCR	Reverse Transcriptase Quantitative Polymerase Chain Reaction
REU	Research Experience for Undergraduates
Sed	Sediment
vls	Virus-like sequence
ZM	Zebra mussel

Abstract:

Globalization and resulting increases in international trade has allowed for the movement of species between almost all ecosystems on Earth. The Great Lakes in particular has seen a number of high impact invasive species that have moved in following the opening of the lakes to intercontinental shipping trade. This research specifically focuses on the benthic community of Lake Michigan which has seen significant community shifts including declines of the amphipods *Diporeia spp*. and the expansion of invasive quagga mussels (*Dreissena rostriformis bugensis*). Little is known about the viral community in the benthos of Lake Michigan and how it may have changed with the establishment of invasive species. I am interested specifically in a group of important, potentially pathogenic viruses, known as Circular Rep Encoding Single Stranded DNA viruses or CRESS-DNA viruses. Here I examined samples from the central Lake Michigan benthos including quagga mussels and sediment cores using metaviromics and qPCR to examine a specific CRESS-DNA virus-like sequence. A specific *Diporeia spp*. associated CRESS-DNA virus (circovirus) known as LM29713 (identified in Hewson et al., 2013) was investigated using qPCR where it was found to be present throughout nearshore and offshore sediment cores. My results suggest that the virus is endemic and has been present in Lake Michigan as long as its host *Diporeia spp.* have been present. I also completed a metavirome analysis of all the CRESS-DNA virus-like sequences found in an offshore sediment core and quagga mussels samples. A comparison of all the different CRESS-DNA virus-like sequences found in the samples using a heat map shows a greater overlap in the viruses found in the quagga mussels of different size and ages and a larger difference between the viruses found in the sediment core layers. Overall the CRESS-DNA virus-like sequences reads

found in each of the libraries were quite different. Still nine specific CRESS-DNA virus-like sequences were found with reads detected in both the deep sediment core layers and the quagga mussels. Cs^{137} dating results of the offshore sediment cores showed two distinct layers, the top 6 cm sediment layer after 1952 and the second before 1952. Both of my sediment core metavirome libraries were from before 1952 indicating that quagga mussel invaders are interacting with CRESS-DNA virus-like sequences likely historically present in the region. A specific CRESS-DNA virus-like sequence, Lake Michigan Quagga Mussel associated CRESS-DNA virus-like sequence 1241 (LMQMvls1241) was found to be abundant in only the quagga mussel metavirome libraries. Hence, I used qPCR to quantify the abundance of the virus-like sequence in quagga mussel samples. Quagga mussel samples from across the central Lake Michigan benthos and Muskegon Lake were compared with samples obtained in the Rybinsk Reservoir in Russia. LMQMvls1241 was not present in samples from the Rybinsk Reservoir in Russia and found only in quagga mussels in the offshore benthos of central Lake Michigan suggesting that it did not originate in the quagga mussels native region (Europe) and are likely part of the Lake Michigan benthic viral consortium. This study shows that quagga mussels are interacting with CRESS-DNA viruses that have been historically present in the Lake Michigan benthos and raises interesting questions about what invertebrate associated CRESS-DNA viruses are infecting.

Chapter 1

Introduction:

Throughout human history the oceans have been a primary source of globalization by means of transportation. Over the last few decades, however, intercontinental trade via shipping has increased by 5.4% per year allowing for a more extreme flux of exotic organisms between ecosystems throughout the world (Keller et al., 2014). Consequences of exotic species movement includes biodiversity loss and native species extinction (Ehrenfeld, 2005). These invasive and non-native species are profoundly changing both terrestrial and aquatic ecosystems worldwide (Keller et al., 2014).

A prime example of ecosystem change from exotic species movement is occurring in the Great Lakes of North America. The Great Lakes have been invaded by more non-native species than any other freshwater ecosystem (Pagnucco et al., 2015). Some of the most problematic invasive species came from the Ponto-Caspian Sea region arriving via ballast water from transoceanic cargo vessels (Vanderploeg, 2002; Alexander, 2009). The most notorious invaders are the dreissenid mussels (beginning with zebra mussels; *Dreissena polymorpha)* first reported in Lake St Clair in 1988 (Nalepa and Schloesser, 1993). Dreissenid mussels are able to proliferate rapidly and thrive throughout the Great Lakes littoral zone due to the abundant plankton food resources, efficient filtering rates selecting for optimum food particulates, and an effective reproductive strategy (Nalepa and Schloesser, 1993; 2014). Following the zebra mussels infestation another dreissenid species, the quagga mussel (*Dreissena bugensis*), arrived in the lakes and has since become even more prolific than the zebra mussels spreading deep into the soft sediments of the offshore benthos of the lower Great Lakes (Karatayev et al., 2015). The

expansion of quagga mussel populations into the offshore benthos and the corresponding declines of the many other species (e.g. *Diporeia spp.)* is transforming the Lake Michigan benthos by cutting the link to higher trophic levels, making the mussel-dominated benthos a major energy sink (Nalepa et al., 2009).

A remaining mystery in Great Lakes benthic ecology is the decline of the amphipod, *Diporeia spp*. (first known as *Pontoporeia spp.). Diporeia spp.* were historically the most abundant taxa burrowing in the soft substrates of the offshore benthos (defined here as greater than 30 m in depth) originally comprising more than 60% of the biomass (Winnell and White, 1984; Sly and Christie, 1992; Nalepa et al., 2006). Beginning in the early 1990s *Diporeia* populations started to decline and now are nearly extirpated from the lower Great Lakes (Barbiero et al., 2011). Several hypotheses have been proposed regarding why *Diporeia* populations crashed including food limitation due to competition with dreissenid mussels (Nalepa et al., 2006; 2009), toxins filtered and concentrated in the psuedofeces (undigested waste products; Dermott et al., 2005) of the mussels, and also potential diseases (Cave and Strychar, 2014), but the definitive cause of the decline has yet to be determined. A study by Hewson et al. (2013a) found that new specific circoviruses (a type of CRESS-DNA virus), including a virus named "LM29173", were in high abundance in *Diporeia spp.* in the then rapidly declining Lake Michigan population suggesting that viruses may have played a role in shifting the Lake Michigan benthic community. However, new research suggests that LM291723 circoviruses did not seem to impart any physiological or gene expression changes in *Diporeia spp.* (Bistolas et al., 2017a; 2017b).

Viruses are biological entities that are obligate parasites and require specific hosts in order to reproduce (Hurst, 2000). Viruses represent one of the greatest sources of biodiversity and are the most abundant biological entities on the planet (Edwards and Rohwer, 2005; Suttle, 2013). Viruses are important in aquatic ecosystems where they play a major role in nutrient cycling and help to control population dynamics of organisms at all trophic levels (Fuhrman, 1999; Breitbart and Rohwer, 2005; Suttle, 2007). New methods in viral metagenomics (i.e. metaviromics), which is the study of viral genetic diversity in the environment without culturing them, has resulted in the discovery of novel viruses including many circular rep-encoding single-stranded DNA viruses (CRESS-DNA viruses) in aquatic environments (Breitbart et al., 2002; Edwards and Rohwer, 2005; Thurber et al., 2009). CRESS-DNA viruses are a group of viruses that include many pathogens of plants, invertebrates, and vertebrates (Rosario et al., 2009). They have been shown to be increasingly prevalent in freshwater environments, however, their role is still not well understood (Rosario et al., 2009; Dayaram et al., 2016). CRESS-DNA viruses have been isolated in ecosystems world-wide from the bottom of the Marianas Trench in the Pacific Ocean (Yoshida et al., 2013), Antarctica lakes (Lopez-Bueno et al., 2009), as well is in the Great Lakes (Hewson et al., 2013a, Bistolas et al., 2017a; 2017b;). CRESS-DNA viruses are speculated to move through insect and other invertebrate populations such as mollusks and mussels in freshwater lakes and are also present in lake benthic surface sediments (Dayaram et al., 2015; 2016). They have also been detected in ballast water samples of transoceanic cargo vessels moving through the Great Lakes region (see Kim et al., 2015).

Viruses adsorb to descending particulates and as they sink into deeper water where they become mixed with anoxic sediments. Through this process, it is believed they are preserved for long periods of time (Hewson and Fuhrman, 2003). In the Black Sea, cyanophages of surface water algae were observed to be present in 7000 year old sediment layers (Coolen, 2011). Several studies have observed CRESS-DNA viruses of invertebrates found in the lake sediments including those of *Daphnia spp.* (Hewson et al., 2013b) and in New Zealand lakes (Dayaram et al., 2016). Hewson et al. (2012) suggests single-stranded DNA viruses are more stable in the abiotic environment (outside of their hosts) than double-stranded DNA viruses and that sediments may serve as environmental reservoirs for single-stranded DNA viral genotypes. Hence, sediment cores in lake ecosystems offer a mechanism to track changes of viral communities over time. Very few if any studies have examined the preservation of invertebrate associated CRESS-DNA viruses deep in sediment cores.

Purpose:

The purpose of this research was to look at a unique part of the nanobiome of one of the currently most abundant benthic organisms (quagga mussels) in Lake Michigan as well as in the benthos (i.e. sediment cores). An often overlooked consequence of invasive species ecology is the transfer of unique members of the microbiome and nanobiome, including potential pathogenic bacteria and viruses of a variety of different organisms. My study offers a unique perspective looking at an emerging important group of viruses, CRESS-DNA viruses, and their prevalence and diversity in the notorious Lake Michigan invader, quagga mussels, as well as their prevalence in the past by looking back in sediment cores. It is the first to document the

CRESS-DNA viruses associated with quagga mussels from parts of both their expanded North American and European ranges.

Scope:

The scope of this research is focused primarily on the central Lake Michigan benthos although I do investigate a small number of samples collected from the Rybinsk Reservoir in Russia. Increased global trade due to globalization is however likely facilitating the movement of CRESS-DNA viruses associated with invasive species to novel ecosystems around the globe.

Assumptions:

This study assumes that all viruses discovered in the metaviromes were found in the mussels specifically those infecting tissues, food sources, symbionts, and those bound to particulate matter that may get caught in the gills and in the sediments and not the waters that may have surrounded the tissues upon preservation. For the quagga mussels contained in the laboratory tanks it was assumed that no viruses were introduced via the food, air, or Milli-Q water supply used. This study assumes that the metaviromes of both the quagga mussels and sediment cores samples are representative of the larger CRESS-DNA virus consortia in the environment they were collected in.

Limitations:

Due to budget and time constraints, this study is limited to only small spatial regions in the central Lake Michigan benthos and the Rybinsk Reservoir in Russia. Spatial and temporal differences in the sediment core and quagga mussel samples are due to unexpected changes and budget in experimental collections and analyses, specifically nearshore and offshore sediment coring and quagga mussel benthic grabs.

Hypothesis:

Null: Quagga mussel invaders did not bring in any unique CRESS-DNA viruses when they invaded the Lakes, but instead are taking in CRESS-DNA viruses via filter feeding that have already been around in the central Lake Michigan benthos.

Alternative: Quagga mussel invaders brought in an associated community of CRESS-DNA viruses when they invaded the Great Lakes.

Research Questions:

1. What does the CRESS-DNA viral consortium look like from quagga mussels and sediment cores collected in the central Lake Michigan benthos?
2. Are the invasive mussels harboring new viruses that were potentially brought in when they invaded the Lake Michigan benthos?
3. Are there specific CRESS DNA viruses that can be found in both the sediment cores and the quagga mussels?

Significance:

This research is significant in that it provides the first look at important and unknown constituency of CRESS-DNA viruses in a pervasive Great Lakes benthic invader, the quagga mussel. The research adds data that will be helpful in understanding the distribution and evolution of CRESS-DNA viruses in benthic invertebrates and aquatic ecosystems across the globe.

References:

Alexander, J., 2009. Pandora's locks: The opening of the Great Lakes-St Lawrence Seaway. East Lansing, MI: Michigan State University Press, 431 p.

Barbiero, R.P., Schmude, K., Lesht, B.M., Riseng, C.M., Warren, G.J., Tuchman, M.L., 2011. Trends in *Diporeia* populations across the Laurentian Great Lakes, 1997-2009. Journal of Great Lakes Research 37, 9-17.

Bistolas, K.S.I., Jackson, E.W., Watkins, J.M., Rudstam, L.G., Hewson, I., 2017a. Distribution of circular single-stranded DNA viruses associated with benthic amphipods of genus *Diporeia* in the Laurentian Great Lakes. Freshwater Biology 62(7), 1220-1231.

Bistolas, K.S.I., Rudstam, L.G., Hewson, I., 2017b. Gene expression of benthic amphipods (genus: *Diporeia*) in relation to a circular ssDNA virus across two Laurentian Great Lakes. PeerJ e3810; DOI 10.7717/peerj.3810.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology 13, 278-284.

Breitbart, M., Salamon, P., Andrees, B., Mahaffy, J., Segal, A., Mead, D., Azam, F., Rohwer, F., 2002. Genomic analysis of uncultured marine viral communities. Proceedings of the National Academy of Sciences U.S.A. 99, 14250-14255.

Brown, J., 2014. Comparison of water column virus populations over a 180+ year time period within the varved sediments of Fayetteville Green Lake. Chapter 4 of Dissertation. Cornell University.

Cave, C.S., Strychar, K. B., 2014. Decline of *Diporeia* in Lake Michigan: Was Disease Associated with Invasive Species the Primary Factor? International Journal of Biology 7, 93.

Coolen, M.J., 2011. 7000 years of *Emiliania huxlyi* viruses in the Black Sea. Science 33, 451-452.

Cook, D.G., Johnson G., 1974. Benthic macroinvertebrates of the St Lawrence Great Lakes. Journal of the Fisheries Research Board of Canada 31, 763-782.

Dayaram, A., Goldstien, S., Arguello-Astroga, G.R., Zawar-Reza, P., Gomez C., Harding, J.S., Varsani, A., 2015. Diverse small circular DNA viruses circulating amongst estuarine molluscs. Infections, Genetics and Evolution 31, 284-295.

Dayaram, A., Galtowitsch, M.L., Arguello-Astorga, G.R., Van Bystervelt, K., Kraberger, S. Stainton, D., Harding J.S, Roumagnac, P., Martin, D.P. Lefeuvre, P., Varsani, A., 2016. Diverse circular replication-associated protein encoding viruses circulating within a lake ecosystem. Infection, Genetics and Evolution 39, 304-316.

Dermott, R., Munawar, M., Bonnell, R., Carrou, S., Nibblock, H., Nalepa, T., Messick, G., 2005. Preliminary investigations for the causes of the disappearance of *Diporeia spp*. from Lake Ontario. Pp. 203-232 in Great Lakes Fisheries Commission technical report. Great Lakes Fisheries Commission. Ann Arbor, MI.

Dunlap, D. S., Ng, T. F. F., Rosario, K., Barbosa, J. G., Greco, A. M., Breitbart, M., Hewson, I., 2013. Molecular and microscopic evidence of viruses in marine copepods. Proceedings of the National Academy of Sciences, 110(4), 1375-1380.

Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. National Review of Microbiology 3, 504-510.

Ehrenfeld, D., 2005. The Environmental Limits of Globalization. Conservation Biology 19(2), 318-326.

Elston, R., 1997. Special topic review: Bivalve mollusk viruses. World Journal of Microbiology and Biotechnology 13, 393-403.

Fitzgerald, S.A., Gardner, W.S., 1993. An algal carbon budget for pelagic-benthic coupling in Lake Michigan. Limnology and Oceanography 38(3), 547-560.

Fuhrman, J. A., 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399(6736), 541-548.

Gastrich, M.D., Leigh-Bell, J.A., Gobler, C.J., Anderson, O.R., Wilhelm, S.W., Bryan, M., 2004. Viruses as potential regulators of regional brown tide blooms caused by the alga, *Aureococcus anophagefferens*. Estuaries 27(1), 112-119.

Hewson, I., Barbosa, J.G., Brown, J.M., Donelan, R.P., Eaglesham, J.B., Eggleston, E.M., LaBarre, B.A., 2012. Temporal dynamics and decay of putatively allochthonous and autochthonous viral genotypes in contrasting freshwater lakes. Applied and Environmental Microbiology 78(18), 6583-6591.

Hewson, I., Eaglesham, J.B., Hook, T.O., LaBarre, B.A., Sepulveda, M.A., Thompson, P.D., Watkins, J.M., Rudstam, L.G., 2013a. Investigation of viruses in *Diporeia spp*. from the Laurentian Great Lakes and Owasco Lake as potential stressors of declining populations. Journal of Great Lakes Research 39, 499-506.

Hewson, I., Fuhrman, J.A., 2003. Viriobenthos production and virioplankton sorptive scavenging by suspended particles in coastal and pelagic waters. Microbiology Ecology 46, 337-347.

Hewson, I., Ng, G., Li, W., LaBarre, B.A., Aguirre, I., Barbosa, J.G., Breitbart, M., Greco, A.W., Kearns, C.M., Looi, Al, Schaffner, L.R., Thompson, P.D., Hairston, N.G., 2013b. Metagenomic identification, season dynamics and potential transmission mechanisms of a Daphnia- associated single- stranded DNA virus in two temperate lakes. Limnology and Oceanography 58, 1605-1620

Hurst, C.J., 2000. Viral Ecology. Academic Press, 659 p.

Karatayev, A.Y., Burlakova, L.E., Padilla, D.K., 2015. Zebra versus quagga mussels: a review of their spread, population dynamics, and ecosystems impacts. Hydrobiologia 746, 97-112.

Keller, R. P., Cadotte, M. W., Sandiford, G., 2014. Invasive Species in a Globalized World. University of Chicago Press, 427 p.

Kim, Y., Aw, G.T., Teal, K.T., Rose, J.B., 2015. Metagenomic Investigation of Viral Communities in Ballast Water. Environmental Science and Technology 49, 8396-8407.

Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., Alcami, A., 2009. High Diversity of the Viral Community from an Antarctic Lake. Science New Series 326, 858-861.

Lymer, D., Logue, J.B., Brussard, C.P.D., Baudoux, A.C., Vrede K., Lindstrom, E.S., 2008. Temporal variation in freshwater viral and bacterial community composition. Freshwater Biology 53, 1163-1175.

Middleboe, M., Jacquet, S., Weinbauer, M., 2008. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. Freshwater Biology 53, 1069-1075.

Murray, A.G., Jackson, G.A., 1992. Viral dynamics: a model of the effects size, shape, motion and abundance of single-celled planktonic organisms and other particles. Marine Ecology Progress Series 89, 103-116.

Nagasaki, K., Tarutani, K., Tomaru, Y., Itakura, S., Tamai, K., Yamanaka, S. Yamaguchi, M., 2002. Possible use of viruses as a microbiological agent against harmful algal blooms. Fisheries Science 68, 497-500.

Nalepa, T.F., Fanslow, D.L., Lang, G.A., 2009. Transformation of the offshore benthic community in Lake Michigan: a recent shift from the native amphipod *Diporeia spp.* to the invasive mussel *Dreissena rostriformis bugensis*. Freshwater Biology 54, 466-479.

Nalepa, T.F., Fanslow, D.L, Iii, A.J.F., Lang, G.A., Eadie, B.J., Quigley, M.A., 2006. Continued disappearance of the benthic amphipod *Diporeia spp.* in Lake Michigan: is there evidence for food limitation? Canadian Journal of Fisheries and Aquatic Science 63, 872-890.

Nalepa T.F., Hartson, D.J., Fanslow, D.L., Lang, G.A., Lozano, S.J., 1998. Declines in benthic macroinvertebrate populations in southern Lake Michigan, 1980-1993. Canadian Journal of Fisheries and Aquatic Science 55, 2402-2413.

Nalepa, T.F., Schloesser, D.W., 1993. Zebra Mussels Biology, Impacts and Control. Lewis Publishers, Boca Raton (810 p.).

Nalepa, T.F., Schloesser, D.W., 2014. Quagga and Zebra Mussels: Biology, Impacts and Control. 2nd edition. CRC Press, Boca Raton (7750 p.).

Ng, T.F.F., Willner, D., Nilsson, C., Lim, Y.W., Schmieder, R., Chau, B., Ruan, Y., Rohwer, F., Breitbart, M., 2010. Vector-based metagenomics for animal virus surveillance. International Journal Infectious Disease 14, E378.

Noble, R.T., and Fuhrman, J.A., 1998. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquatic Microbial Ecology 14, 113-118.

Pagnucco, K.S., Maynard, G.A., Fera, S.A., Yan, N.D. Nalepa, T.F., Ricciardi, A., 2015. The future of species invasions in the Great Lakes-St. Lawrence River basin. Journal of Great Lakes Research 41, 96-107.

Rosario, K., Duffy, S., Breitbart, M., 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. Journal of General Virology 90, 2418-2424.

Sly, P.G., Christie, W.J., 1992. Factors influencing densities and distributions of *Pontoporeia hoyi* in Lake Ontario. Hydrobiologia 235, 321-352.

Suttle, C.A., 2007. Marine viruses- major players in the global ecosystem. National Review of Microbiology 5, 801-812.

Suttle, C.A., 2013. Viruses: unlocking the greatest biodiversity on Earth. Genome 10, 542-544.

Thingstad, T.F., 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnology and Oceanography 45, 1320-1328.

Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. Nature protocols 4(4), 470-483.

Van Etten, J.L., Van Etten, C.H., Johnson, J.K., Burbank D.E., 1985. A survey for viruses from freshwater that infect a eukaryotic Chlorella-like green alga. Applied Environment Microbiology 49, 1326-1328.

Vanderploeg, H.A. Liebid, J.R., Nalepa, T.F., Fahnenstiel, G.L., Pothoven, S.A., 2010. Dreissena and the disappearance of the spring phytoplankton bloom in Lake Michigan. Journal of Great Lakes Research 36, 50-59.

Vanderploeg, H.A, Nalepa, T.F., Jude, D.J., Mills, E.L., Holeck, K.T., Liebig, J.R., Grigorovich, I.A., Ojaveer, H., 2002. Dispersal and emerging ecological impacts of Ponto- Caspian species in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 59, 1209-1228.

Watkins, J.M., Dermott, R., Lozano, S.J., Mills, E.L. Rudstam L.G., Scharold, J.V., 2007. Evidence for remote effects of dreissenid mussels on the amphipod Diporeia: analysis of Lake Ontario benthic surveys, 1972-2003. Journal of Great Lakes Research 33, 642-657.

Wilhlem, S.W., Suttle, C.A., 1999. Viruses and nutrient cycles in the sea. Bioscience 49, 781-788.

Winnell, M.H., White, D.S., 1984. Ecology of Shallow and Deep Water Populations of Pontoporiea hoyi (Smith) (Amphipoda) in Lake Michigan. Freshwater Invertebrate Ecology 3, 118-138.

Winters, A.D., Fitzgerald S., Brendan, T.O., Napela, T., Faisal, M., 2014. Spatio-temporal dynamics of parasites infecting *Diporeia spp.* (Amphipoda, Gammaridae) in southern Lake Michigan (USA). Journal of Invertebrate Pathology 121, 37-35.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., Takai, K., 2013. Metagenomic Analysis of Viral Communites in (Hado) Pelagic Seidments. PLoS One 8(2), e5727.

Chapter 2 (Manuscript- Journal of Great Lakes Research)

Surveying the Circular Rep Encoding Single Stranded (CRESS) DNA viral consortium in quagga mussels (*Dreissena rostriformis bugensis*) and sediments from the central Lake Michigan benthos

Gezon, N.R.[1], Bryant, R.[2], Haywick, D.[3], Sanders, J.[3], Hewson, I.[4], & Strychar, K.B.[1*]

[1] Annis Water Resources Institute, Grand Valley State University, 740 W. Shoreline Dr. Muskegon, MI 49441-1678

[2]Rice University 6100 Main St, Houston, TX 77005

[3]University of South Alabama 307 N. University Blvd. #130, Mobile, AL 36688

[4] Department of Microbiology, Cornell University, 403 Wing Hall, Ithaca NY 14850

*Corresponding Author

Email: strychak@gvsu.edu

Phone: (616) 331-8796

Abstract:

Circular Rep Encoding Single Stranded (CRESS) DNA viruses have come into the spotlight as an important group of viruses associated with invertebrates in freshwater aquatic ecosystems. Recent work in the Great Lakes has uncovered interesting information on the ecology and diversity of CRESS-DNA viruses associated with the amphipods, *Diporeia spp.* The Lake Michigan benthos has changed considerably with drastic declines of populations of *Diporeia spp.* and concurrent increase in numbers of the invasive quagga mussel (*Dreissena rostriformis bugensis*). The CRESS-DNA viruses that are associated with the invasive quagga mussels are, however, unknown. The purpose of this study was to characterize the CRESS-DNA viral constituency found in both invasive quagga mussels and in sediment cores collected in the central Lake Michigan benthos, and also look for a specific CRESS-DNA virus-like sequence in quagga mussels from their European expanded range with samples collected in the Rybinsk Reservoir in Russia. Metavirome libraries of CRESS-DNA viruses were completed for two size classes (large >25mm SL and small < 15mm SL) of Lake Michigan quagga mussels and for two different sediment core layers. A specific virus-like sequence, LMQMvls1241, identified in the metaviromes was further investigated using targeted qPCR and RT qPCR. Results show strong differences in the viral constituencies of the quagga mussels and sediment cores with some overlap. A particular virus-like sequence, LMQMvls8177, was identified in the large quagga mussel metavirome library and was also found to be highly prevalent deep in sediment core layers. Cs^{137} radioactive dating results show LMQMvls8177 to be present in the sediment cores before the arrival of quagga mussels suggesting that the new invaders are interacting with CRESS-DNA viruses historically present in the Lake Michigan benthos. Targeted qPCR results

showed LMQMvls1241 virus-like sequence to be widespread in the offshore benthos of Lake Michigan while not present in samples collected in Muskegon Lake or in the Rybinsk Reservoir in Russia although RT qPCR of LMQMvls1241 showed no active replication of the virus in the quagga mussel samples. I speculate that some of the CRESS-DNA virus-like sequences found in the quagga mussel libraries could be associated with a single celled host organism either present in or on quagga mussels as symbionts or present in Lake Michigan that the quagga mussels are picking up via filter feeding. Overall results help reveal more about the ecology of CRESS-DNA viruses in freshwater aquatic ecosystems.

Introduction:

The benthic community in the Great Lakes has shifted considerably in the past decades with reductions in nutrient loading following the passing of the Clean Water Act (1972), the establishment of nonindigenous species including dreissenid mussels, and displacement and rapid declines of populations of the native amphipod *Diporeia spp.* (Nalepa et al., 1998; 2009; Barbiero et al., 2011). Dreissenid mussels, which includes both zebra mussels *(Dreissena polymorpha)* and quagga mussels *(Dreissena rostriformis bugensis*) are native to the Ponto-Caspian Sea region and Dnieper River region in Eastern Europe (Nalepa and Schloesser, 2014). It has been speculated that both species arrived in the Great Lakes via ballast water associated with transoceanic cargo ships (Nalepa and Schloesser, 1993; 2014). Zebra mussels became predominately established in hard substrates of the littoral zone and nearshore waters (<30m of water depth) followed by quagga mussels which have spread into the deep soft sediment substrates of the offshore benthos (Nalepa and Schloesser, 1993; 2014). In Lake Michigan today, zebra mussel populations have declined and quagga mussels now dominate (Cuhiler and Aguliar, 2013). Dreissenid mussel adults are sessile, while the larva are planktonic (known as veligers) helping them to disperse throughout the lakes using natural currents (Nalepa and Scholesser, 1993; 2014). The dreissenid mussel populations have large filter feeding capabilities that have also altered plankton communities, moving biomass and carbon into the benthos (Vanderploeg et al., 2002). In Lake Michigan in particular, the establishment of large populations of dreissenid mussels has altered many populations of native species both in the plankton (e.g. algae and zooplankton) and in the benthos (*Dipoeria spp.*; Cuhiler and Aguliar, 2013; Karatayev et al., 2015).

Diporeia spp. are a glacial relic benthic amphipod species that was common to soft substrates of the offshore waters of Lake Michigan (Marzolf, 1965). *Diporeia* was once widespread in the Lake Michigan benthos but has since been mostly extirpated across their range (Barbiero et al., 2011). Even amongst a substantial number of studies (Nalepa et al., 2006; Watkins et al., 2007; Nalepa et al., 2009; Barbiero et al., 2011; Watkins et al., 2012) the precise mechanisms and causation of the declines remains a mystery. Although *Diporeia* population declines have been correlated with the large increases in dreissenid populations no direct causative link has been established (Nalepa et al., 2006). Two studies (Nalepa et al., 2009; Watkins et al., 2012), however, have shown *Diporeia* declines occurring prior to the establishment of dreissenid mussels in Lake Michigan and in Lake Ontario and the two populations have been shown to coexist in the Finger Lakes region of NY (Watkins et al., 2012). Nalepa et al. (2009) conducted a field survey and a mesocosm experiment but found no evidence supporting indirect competition over spring diatom bloom food resources. Similarly, Watkins et al. (2012) observed no trends with *Diporeia* populations in Lake Ontario and benthic carbon flux, lake productivity and fish predation impacts. Research into potential associated pathogens, eukaryotic parasites, bacteria, and viruses was unable to explain variation in *Diporeia* population declines (Hewson et al., 2013a; Winters et al., 2014; Bistolas et al., 2017). A specific group of viruses, Circular Replication Initiator Protein encoding single stranded DNA viruses (CRESS-DNA viruses) were found to be in high abundance in the declining Lake Michigan populations of *Diporeia spp.* (Hewson et al. 2013a). Recent extensive research into *Diporeia* associated CRESS-DNA viruses by Bistolas et al. (2017a; 2017b) reveals that recurrent and widespread CRESS-DNA viral genotypes are found in *Diporeia* species across the Great Lakes but

that a specifically identified CRESS-DNA viral genotype, LM29173 (first identified in Hewson et al., 2013a), did not impart any physiological detriment to *Diporeia* in virus inoculation experiments (Bistolas et al., 2017a) and that no gene expression pathways were initiated in *Diporeia* found in the field with high LM29173 viral load (Bistolas et al., 2017b).

In general, viruses are the most abundant and diverse biological entities on Earth where they play a major role in nutrient cycling (e.g. bacteriophages in ocean carbon cycling; Fuhrman, 1999) and host species population dynamics (Breitbart and Rohwer, 2005; Edwards and Rohwer, 2005; Suttle, 2007). Increases in the discovery of diversity and abundances of viruses in aquatic environments have been brought on by using the method of viral metagenomics (metaviromics; Thurber et al., 2009). Metaviromics, which uses shotgun sequencing of purified viral particles to identify all the known viral genetic components of samples without having to culture the viruses, has been an effective tool for characterizing CRESS-DNA viral abundance and diversity in a range of different aquatic environs including freshwater lakes (Thurber et al., 2009; Rosario and Breitbart, 2011; Roux et al., 2012; Hewson et al., 2013; Dayaram et al., 2016; Bistolas et al., 2017). CRESS-DNA viruses are a specific group of single stranded DNA viruses, some of the smallest known to infect eukaryotic organisms, which encode a replication initiator protein that is well conserved in members of this viral taxonomy (Rosario et al., 2012). CRESS-DNA viruses have been found to be a very common component of invertebrate nanobiomes but the exact role the viruses play remains poorly described (Rosario and Breitbart, 2011). CRESS-DNA viruses are found to be associated with planktonic invertebrates such as copepods (Dunlap et al., 2013) and *Daphnia spp.* where in one particular study, CRESS-DNA viral load maxima were shown to occur before major *Daphnia* population declines in a freshwater lake (Hewson

et al., 2013b). Research in freshwater lakes in New Zealand characterized a variety of CRESS-DNA viruses found in mollusk and gastropod species in the littoral zone (Dayaram et al., 2013; 2016). CRESS-DNA viruses have also been detected in lake sediments (Dayaram et al., 2016). In rich organic sediments virus-like particles have a slower decay rate because of the inhibition of extracellular nucleases due to chelating properties in the organic compounds and clay materials (Lorenz et al., 1981). Dynamics of CRESS-DNA viruses in an extensively ecologically disturbed freshwater benthic ecosystem have not been investigated to date.

In this study I focused on the CRESS-DNA viruses found in the central Lake Michigan benthos. My goal was to compare and contrast between current and prior known CRESS-DNA viruses and to determine whether either type of virus (e.g. current vs. prior known viral types) exists in the newly dominant quagga mussels. My objectives were: (1) to use metaviromics to describe the CRESS-DNA virus-like sequences found in the invasive quagga mussels in the offshore benthos of central Lake Michigan in a location where *Diporeia* species were historically dominate, (2) to establish a historical record of the CRESS-DNA virus-like sequences found in the central Lake Michigan by characterizing the CRESS-DNA virus-like sequences found in offshore sediment cores samples, specifically focused on the deeper layers before quagga mussels invaders established in the benthos, (3) to compare and contrast the CRESS-DNA virus-like sequences found in the quagga mussels and sediment cores and also with those identified in previous studies of CRESS-DNA viruses found in *Diporeia spp.* in the Lake Michigan benthos, and (4) to examine the distribution of a CRESS-DNA virus-like sequences identified in quagga mussels metavirome libraries to describe its distribution in quagga mussel populations in the central Lake Michigan benthos, Muskegon Lake, and also in the European expanded range of

the quagga mussels in the Rybinsk Reservoir in Russia. I hypothesized that quagga mussels may have brought in a unique consortia of associated CRESS-DNA viruses when they invaded the Great Lakes.

Methods:

Quagga mussel collection and preservation

Initial quagga mussels were collected in approximately 47 meters of water in Lake Michigan west of Holland, MI (Holland Deep, Figure 1A) on June 23, 2016 using a petite ponar grab (WILDCO). This depth and specific location were chosen based on (1) where *Diporeia* species were historically observed to exist in high abundance (Winnell and White 1984) and (2) in close proximity to a previously documented NOAA benthic sampling site. Mussels were sorted on deck of a boat by shell length into three arbitrarily assigned size categories: small (<15mm), medium (15-25mm), and large (>25mm) (Table 1). Shell length was measured in millimeters from the base of the hinge ligament to tip of the shell using a caliper. Approximately 30 mussels of each size class were stored in containers containing 95% ethanol for long term storage. Any remaining quagga mussels were maintained in water collected from the site of origin to be stored in the lab for future use (see chapter 3 extended methodology for more detail). The quagga mussels sampled were found to be the profunda morphotype based on visual assessment of rounded ventral edges, weak pigmentation, and elongated siphon (Dermott and Munawar, 1993). Five large and five small quagga mussel samples from 47 m were used in the metavirome analysis (Holland Deep, Figure 1A). Additional quagga mussels were collected in the central benthos of Lake Michigan and in Muskegon Lake in the early fall of

2016 using a WILDCO ponar grab (LM Sites 1-5 & ML Sites 6-7; Figure 1A). These mussel samples were similarly sorted by size on the boat deck but then immediately frozen in liquid nitrogen (-196°C) for transport and then stored at -80°C in a lab at AWRI in Muskegon MI. Visual assessment of these quagga mussels showed that the mussels collected from the offshore Lake Michigan water were also of the profunda morphotype while the Muskegon lake samples were the epilimnotic morphotype (dark pigmentation, squat fat appearance; Dermott and Munawar, 1993). Quagga mussels from overseas came from the Rybinsk Reservoir along the Volga River in Russia where they were collected in 6-8m of water (Dr. Vera Pavlova pers. comm.; Figure 1B) using a ponar grab in July 2017, stored in 95% ethanol, and couriered to our lab located in Muskegon, MI. The ethanol had to be drained during international boarder inspections but was immediately recharged upon arrival at our institution. Visual inspection of the quagga mussels showed that these samples were the epilimnotic morphotype (dark pigmentation, squat fat appearance; Dermott and Munawar, 1993)

Nearshore sediment cores

Six sediment cores samples were collected using an industrial SDI vibracore system on a self-propelled freight barge off the port of Muskegon (Alpine; White Lake Dock and Dredge, Whitehall, MI) on July 19, 2016. Sediment core and quagga mussel sampling took place in different locations due to project logistics and budget limitations. Three cores were taken near the NOAA benthic survey sampling site M29 in approximately 29 m of water (Nearshore Cores 1-2; Figure 1A). Each core was extracted in a polycarbonate tube of 60 cm in length attached to the vibrator coring unit (VCU, Alpine) which uses high frequency vibration to dig the sediment core into the bottom of the lake. Using an electric winch system, the empty core was attached

to a steel cable and lowered to the benthos. Once the VCU touched bottom, vibration was applied to inject the empty tube into the sediment to collect the core. Upon completion, the sediment core was slowly removed from the benthic substrate and brought back to the surface. A plastic one-way valve cap (core catcher) on the end of the tube prevents the sediment from falling out as it is brought up to the surface. When the core was brought to the surface it was immediately capped, taped shut, labeled and placed in an Igloo cooler with dry ice (-78.5°C) in order to minimize light exposure of the cores. The sediment core sampling was repeated for each core taken. Upon returning to the lab at AWRI, all sediment cores were immediately frozen in a -80 °C freezer to prevent viral DNA degradation.

Offshore sediment cores

Offshore sediment cores were taken aboard the W.G. Jackson research vessel (AWRI) using a gravity coring system (WILDCO) on July 21, 2016 (Offshore Core 3; Figure 1). The VCU used to collect nearshore sediment cores was not used offshore as it did not have the capability to take samples in deeper waters due to lack of cable length. Four sediment cores were collected offshore in depths ranging from 105 m to 120 m. The deep-water gravity corer (DGC) was lowered to the benthic substrate on a steel cable winch followed by a messenger weight sent to the bottom to trigger a vacuum mechanism in the coring device to secure the sediment core in the tube. For each sediment core collected a polycarbonate tube approximately 1 m in length was attached to a weighted head unit. The DGC consisted of a locking mechanism and end cap that allowed sediment to enter but prevented any loss (i.e. a one way valve). When the core was brought to the surface light exposure was minimized and the core was quickly

placed in coolers filled with dry ice (-78.5°C). Upon returning to the lab at AWRI, the cores were immediately frozen in a -80°C freezer.

Sediment core subsampling procedure

To compare differences in the nearshore and offshore sediment composition, three sediment cores, nearshore core 1, nearshore core 2, and offshore core 3 were subsampled (Figure 1A). Only offshore sediment core 3 was used for metaviromic analysis because the initial evaluation of high organic content (silt) made it the most likely core to have good viral deposition and preservation. The core samples were thawed at room temperature for approximately 15-30 minutes before subsampling. A PerforMax band saw (22.86cm) with a Kent diamond coated blade was then used to cut the frozen cores lengthwise, followed by measured lengths of each core and photographic documentation. One half of each sediment core was used for further sampling and the second half was returned to the -80°C freezer for long term storage. The half used for further sub-sampling was dissected into eight approximately equal subsections based on the total core length. Each subsampled core section was labeled subsample 1-8 (later referred to as SS1-SS8; Figure 2) and frozen at -80°C until further analyses could be conducted. This process took approximately one hour to complete per sediment core.

The sediment core subsamples used for genomic analyses were removed from the -80°C freezer and transported in an Igloo cooler containing dry ice (-78.5°C) for 24 hours to allow for transportation to Ithaca, NY. Sediment core subsections were thawed for approximately one hour prior to continuation of subsampling. For each core subsection used for genomic analyses,

a sterile plastic spoon (sterilized with 95% ethanol) was used to scrape away the surface sediments on the top of the core section that may have been contaminated during subsampling. Sections were then measured and then subsampled twice in the center of the core using new sterile plastic spoons (Figure 2). The other half of each core was kept for Cesium-137 (Cs^{137}) isotopic age-dating and grain size analysis in laboratories located at the University of South Alabama. Cesium-137 content was determined by measuring gamma ray emissions at specific frequencies using a Princeton Gamma-Tech N-type high purity intrinsic germanium detector. Signals were processed using a Princeton Gamma-Tech System 8000 multichannel analyzer and further analyzed using Princeton Gamma-Tech Quantum GeD software. Grain size analysis was conducted using the pipette and sieve method (Coventry and Fett, 1979; Haywick, 2006) using 47 mm diameter glass fiber filters (pore diameter 0.5µm) to determine silt and clay fractions and sieving to determine sand and gravel content.

Quagga mussel and sediment core metavirome preparation and analysis

Metavirome preparation followed the methods of Hewson et al. (2013a) with modifications. For quagga mussels, two libraries were created, one for large quagga mussels with shell lengths greater than 25mm and one for small quagga mussels with shell lengths less than 15mm. For each quagga mussel library, five individual mussels (n=5) were used. The internal tissues were removed and homogenized together mechanically with sterile forceps. For benthic sediments, four sediment core subsamples with varying mass (1.27 g from SS2, 4.55 g from SS4, 2.81 from SS6, and 1.54 from SS8; Table 1) were taken from the offshore core 3 (Figure 1A). Quagga mussel and sediment samples used for the metavirome library preparations (n=6) were individually (in separate containers) homogenized in 10mL virus free

(0.02 µm filtered) phosphate buffered saline (PBS) solution. The mussel homogenate was filtered using a 0.2 µm polyethersulfone (PES) syringe filters (Sigma Aldrich, Saint Louis, MO, USA) and then Polyethylen glycol (PEG) -8000 (Sigma Aldrich, Saint Louis, MO,USA) was added and the viruses were precipitated overnight at 4°C. Following overnight precipitation, the mussel homogenate samples were ultra-centrifuged at 15,000 x g and the supernatant discarded. Pellets containing the viruses were resuspended in 1 mL of filtered nuclease free water and then treated with RNAse A (ThermoFisher Scientifc, Waltham, MA, USA) and DNAse 1 (ThermoFisher Scientific, Waltham, MA, USA) at 37°C for 2 hours in order to remove all nucleic acids that were not protected by capsids (i.e. nucleic acids not viral in origin). Following the incubation treatment viral DNA was extracted from the viruses using a ZR DNA tissue kit (Zymo Research, Irvine, CA, USA). The extracted viral DNA was then amplified using a Genomiphi V2 multiple displacement amplification kit (Sigma Aldrich, Saint Louis, MO, USA). It should be noted that Genomiphi V2 multiple displacement was used because I was specifically interested in examining the CRESS-DNA viruses found in these samples and this method has a known bias towards CRESS-DNA viruses in metavirome preparations (Kim and Bae, 2011). Five micrograms of amplified DNA from each of the samples were sequenced using the Illumina platform at Cornell University Core Laboratories Center (Ithaca, NY, USA).

Following sequencing, the files were trimmed and analyzed in CLC Genomics Workbench (v4.0 Qiagen, Hilden, Germany). Sequence libraries were trimmed for ambiguous nucleotides and sequences bigger and smaller than the expected 250 bp. They were then assembled into contiguous sequences using 0.8 identity and 0.5 overlap in CLC Genomics Workbench (v4.0 Qiagen, Hilden, Germany). Following assembly, the tBLASTx algorithm was used against an in-

house database (Hewson Lab) of CRESS-DNA viral genomes accumulated from NCBI. Any contig matches were then filtered further using the BLASTx algorithm against the non-redundant database at NCBI to determine if the sequence hits resembled prokaryotic, non-human eukaryotic or other viral (bacteriophages) sequences using an e-value cut-off of 0.001. All contigs that resembled any prokaryotic, non-human eukaryotic or other viral sequences were discarded from further analyses because I was interested in only potential eukaryotic CRESS-DNA virus- like sequences. A final dataset was then created in MS excel (2016) with each of the unique CRESS-DNA virus-like sequence contigs and the number of iterations (reads) the identified CRESS-DNA virus-like sequences appeared in each sample library. Heat maps were created in Program R version 3.3.2 (R Core Development Team, 2017) in order to visual the differences in the virus-like sequence reads found in each library. One heat map was created to take into account the number of reads relative to library size by dividing the number of reads of each virus-like sequence by the total size of the library in which is what it was found. Another heat map was created to take into account the number of reads relative to total contigs found in each library by dividing the number of reads of each virus-like sequence by the total contigs identified in each library. Additionally, nonmetric dimensional scaling (NMDS; Clarke, 1993) was used to visual the similarity and dissimilarity between the CRESS-DNA virus-like sequence read recruitment across the different libraries but these results were not reported and discussed (see supplementary material for methods and results; Figure S2).

Sediment and Quagga Mussel DNA extraction

Sediment DNA from the different core subsamples was extracted following the procedure described in ZR Soil MiniPrep Kits (Zymo Research, Irvine, CA, USA). Approximately

25 mg of sediment was used for each of the extractions. Two extractions were completed for each of the eight different sediment core subsamples in each of the three sediment cores (n=48). Quagga mussel DNA was similarly extracted following the procedure described by ZR Tissue and Insect MiniPrep Kits (Zymo Research, Irvine, CA, USA). Prior to mussel DNA extractions, however, samples were first removed from a 95% ethanol solution or thawed from -80°C for 15 minutes and the total shell length (mm) was measured using a caliper along with the total wet mass (mg) with and without the shells. Quagga mussel extractions consisted of using approximately 25 mg of homogenized internal tissue. The tissues were removed from the mussel shell and homogenized with sterilized forceps to ensure a mixture of all the different tissue types for extraction.

Lake Michigan Quagga Mussel CRESS-DNA virus-like sequence contig1241 (LMQMvls1241) qPCR

A specific CRESS-DNA virus-like sequence, LMQMvls1241, was identified in the quagga mussel metavirome library and chosen for qPCR quantification. This specific virus-like sequence was analyzed using BLASTx and found to align most similarly with the rep-gene of a Bat circovirus (61% identity; Table 3). It was chosen for qPCR quantification because it was strongly represented (based on reads) and present in only the quagga mussel (both small <15mm SL and large >25mm SL sized mussels) libraries and not the sediment layer libraries. Primers and probes specific to an open reading frame (ORF) region (a nucleotide sequence that codes for a protein) contig1241 sequence, which was found using NCBI's ORFfinder, were designed using the program PRIMER3 (Rozen and Skaletsky, 2000; Table 5). A 100 bp oligonucleotide standard (Table 5) was created from the contig 1241 CRESS-DNA virus-like sequence based on the BLASTx alignments and ORFfinder results (NCBI). The primers were tested around this oligonucleotide

standard using polymerase chain reaction (PCR) and products were visualized using gel electrophoresis (see extended methodology for more details). Samples analyzed include quagga mussels collected at the Holland Deep location (Figure 1A; n=6, 2 for each of the 3 size classes; small <15 mm SL, medium 15 mm to 25 mm SL, and large 25 mm SL), quagga mussels collected at the Holland deep location and kept alive in lab tanks for 6 months (n=3 for each of the 3 size classes), those collected from central offshore Lake Michigan and Muskegon Lake in fall 2016 (Figure 1A; n=17; 3 mussels from each of the site with the exception of site 2 and site 6 in which only 1 mussel was used because of small samples), and those collected in the Rybinsk Reservoir in Russia in summer 2017 (Figure 1B; n=9, 3 for each of the 3 size classes; small <15 mm SL, medium 15 mm to 25 mm SL, and large 25mm SL). DNA was extracted using ZR Tissue & Insect Tissue MicroPrep extraction kits (Zymo Research, Irvine, CA, USA). DNA was also extracted from sediment collected at the Holland Deep (Figure1) location using ZR Soil MicroPrep extraction kits (n=1). Quantitative PCR (qPCR) was run on an Applied Biosystems Step One Plus qPCR machine following the methods of Hewson et al. 2013a with slight modifications. Standards, primers, and probes used along with thermocycling conditions are listed in Table S1. Briefly, each sample was run in triplicate and one experimental replicate was tested for inhibition by spiking the sample with one microliter of 10^4 standard sequence (the 100 bp oligo sequence that represents the LMQMvls1241).

In order to look for active replication of LMQMvls1241, reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) was run on quagga mussels collected during the fall from a central Lake Michigan transect (LM Sites 1-5; Figure 1) and maintained in our lab tanks for six months, which were collected in the same location where the quagga mussels used

for the metavirome were collected (Holland Deep, Figure 1). Holland Deep samples preserved in 95% ethanol that were used to construct the metavirome libraries were not used because RNA degrades over time in ethanol. RNA was extracted from the quagga mussel samples, which were previously flash frozen in liquid nitrogen and stored in a -80°C freezer, after they had been thawed on ice using a ZR Tissue & Insect RNR MicroPrep extraction kit (Zymo Research, Irvine, CA, USA). The extracted RNA was inspected using a nanodrop, transformed to cDNA in a one-step synthesis reaction using Invitrogen VILO cDNA synthesis kit, and then stored in the -20°C freezer for qPCR analysis (Invitrogen, ThermoScientific Waltham, MA, USA).

To compare LMQMvls1241 qPCR data each measurement was converted to viral load which is the number of LMQMvls1241 viral rep gene copies detected per mg of tissue or sediment. The average LMQMvls1241 viral load value of each the three qPCR technical replicates was used for analyses. To test for statistical differences between the measured LMQMvls1241 viral loads across the different sites, a Kruskal Wallis test was run with a Mann-Whitney multiple comparison tests (nonparametric because the data was not normally distributed). All analyses were run in the program R version 3.3.2 (R Core Development Team 2017).

Results:

Lake Michigan quagga mussels (Dreissena bugensis) and sediment cores metavirome analysis

Creation of metaviromes was successful for two of the four sediment core layers (SS2 and SS4; Figure 3) and both large (>25mm shell length) and small (< 15mm shell length) quagga mussel samples. The metavirome libraries ranged in size from ~1.7 to 3.5 million trimmed reads

(Table 1). The trimmed reads were assembled into contiguous sequences with 3076 contigs for small quagga mussels (< 15 mm shell length), 8417 contigs for large quagga mussels(> 25 mm shell length), 1081 contigs for sediment core subsample 2 (-136 mm deep) and 3629 contigs for sediment core subsample 4 (-290 mm deep; Table 1). A combined total of 54 novel CRESS-DNA virus-like sequences (<70% similarity to known genotypes; Bistolas et al., 2017a) were identified in the samples based on BLASTx results (NCBI; Table 1). Thirty one CRESS-DNA virus-like sequences, with reference lengths ranging from 224nt to 1911nt were identified in the large quagga mussel metavirome library (Table 1). Seven CRESS-DNA virus-like sequences with reference lengths ranging from 311nt to 932nt, were identified in the small quagga mussel metavirome library (Table 1). Thirteen CRESS-DNA virus- like sequences, with references lengths ranging from 348nt to 12074nt, were identified in the sediment layer SS2 library (Table 1). Three CRESS-DNA virus-like sequences, with references lengths ranging from 380nt to 2983nt, were identified in the sediment layer SS4 library (Table 1). Reads for the different CRESS-DNA virus-like sequences identified overlapped across the sample libraries (Figure 3).The small quagga mussel metavirome had reads for 23 of the CRESS-DNA virus-like sequences identified in all the libraries, the large quagga mussel metavirome contained reads for 52 of the CRESS-DNA virus-like sequences identified across all the libraries and both sediment core subsamples contained reads for 16 CRESS-DNA virus-like sequences detected across all the libraries (Table 1). None of the CRESS-DNA virus-like sequences identified had elements of a complete viral genome and therefore no CRESS-DNA viral genomes were completed in this study. The sequences for LM29173 *Diporeia*-associated circoviruses (identified in Hewson et al. 2013) were not detected in any of the metavirome libraries. Visualization of the CRESS-DNA

virus-like sequence consortia in the different libraries using a heat map showed greater overlap in the CRESS-DNA virus-like sequences detected between the different sized quagga mussels than between the different offshore sediment core layers but overall the CRESS-DNA viral consortia found in each library were quite different from one another (Figure 3).

Within the metavirome libraries, five interesting CRESS-DNA virus-like sequences were selected for further investigation because they had BLASTx results most similar to the replicator initiator protein genes of previously identified CRESS-DNA viruses (Table 2). The first virus-like sequence, which I named Lake Michigan Sediment Subsample 2 virus-like sequence 625 (LMSS2vls625) was identified in the offshore sediment core SS2 layer library with a reference length of 1934nt with the first BLASTx hit being the Rep domain protein [*Phaeocystis globosa* virus] (e-value= 5e-70;Table 2). The next three virus-like sequences were identified in the large quagga mussel library: (1) LMQMLvls3293 with a reference length of 1911nt and first BLASTx hit being hypothetical protein [uncultured marine virus], (2) LMQMLvls3301 with a reference length of 1397nt and first BLASTx hit being hypothetical protein [uncultured marine virus], and (3) LMQMLvls8177 with a reference length of 488nt and first BLASTx hit being Rep domain protein [*Phaeocystis globosa* virus](e-values: 2e-32, 4e-90,1e-43 ; Table 3). The final virus-like sequence of interest, LMQMSvls1241, was identified in the small quagga mussel library with a reference length of 596nt and first BLASTx hit being Rep [Bat circovirus] (e-value 1e-57; Table 2). Nine CRESS-DNA virus-like sequences had reads in both the quagga mussel and sediment core samples with read overlap and BLASTx results (see Table 3).

LMQMvls1241 qPCR analysis

Quantitative PCR (qPCR) analysis of LMQMvls1241 in the quagga mussel samples showed the virus-like sequence was present in all the offshore Lake Michigan samples but not in those collected in the Muskegon Lake and the Rybinsk Reservoir in Russia. There was no significant difference in the LMQMvls1241 viral load (copies per milligram tissue) between the different sized (small < 15 mm SL, medium 15 mm-25 mm SL, and large >25 mm SL) quagga mussels collected at the Holland Deep location (n= 6, Kruskal Wallis test, p=0.32). All qPCR measurements of the nine quagga mussel samples (of varying sizes) from the Rybinsk Reservoir in Russia and the four quagga mussel samples from Muskegon Lake showed a C_t value that appeared at or lower to the value in the no template control runs indicating that LMQMvls1241 sequence was below the threshold of detection and likely not present in the samples. All 13 of the offshore Lake Michigan quagga mussels (fall 2016) investigated tested positive for the virus-like sequence with the average viral load (gene copies per mg tissue) calculated to be 401 copies per mg quagga mussel tissue (n=13, 401.14 $^{+}/_{-}$ 393.965; Figures 4 and 8). In the offshore Lake Michigan samples, individual mussels collected at sites 4 and 5 had the highest viral loads measured (site 4= 1028 copies per milligram tissue; site 5= 1419 copies per milligram tissue). The highest viral load in an individual mussel was 1419 copies per mg tissue found in a mussel collected at offshore Lake Michigan site 5 while the lowest viral load was 93 copies per mg found in a mussel collected at site 3 (Figure 4). Overall there was a significant difference between the average viral load measured in the mussels at the different sites in Lake Michigan (Kruskal Wallis test p=0.0017; Figure 4) with the significant difference being between Sites 1

and 4 and Sites 3 and 4 (Wilcoxon Pairwise Comparison, Holm correction; p= 0.025 and p=0.005 respectively). Similarly there was a statistically significant difference between the average viral load measured in all the Lake Michigan sites compared to all those measured in Muskegon Lake and the Rybinsk Reservoir in Russia (with sites 2 and sites 7 removed because n= 1; Kruskal Wallis Test p=0.007).

Reverse Transcriptase (RT) qPCR analysis showed no active replication of LMQMvls1241 sequences in the tissues of Lake Michigan quagga mussels collected during the fall 2016 sampling. All qPCR measurements showed a C_t value that appeared at or lower to the value measured in the no template control runs indicating that no LMQMvls1241 cDNA was below the threshold of detection and likely not present in the samples.

Sediment core grain size analysis and Cs^{137} dating

Grain size analysis of sediment within the cores reveals information about sedimentary conditions, particularly the energy of deposition. The two nearshore cores extracted in this study were both dominated by sand particles (2 to 0.063 mm) with lesser amounts of silt (0.063 to 0.004 mm) and virtually no clay (<0.004 mm). With the exception of the top 10 cm which contained less than 5% gravel (>2 mm) and very coarse sand (vc: 2 to 1 mm), nearshore core 1 was primarily composed of very fine sand (vf: 0.125 to 0.063 mm; Figure 5). Nearshore core 2 was more stratified. The top 20 cm was primarily composed of coarse (c: 1 to 0.5 mm) and medium (m: 0.5 to 0.25 mm) sand (Figure 6). The remainder of the core was mostly vf sand. Given the domination by sand sized particles in the two nearshore cores, we conclude that both sites were subject to relatively energetic wave/current conditions. The offshore core 3 was

dominated by silt and clay. Very fine sand only made up 5 to 15% of the entire core confirming a lower energy of deposition (Figure 7).

In order to date the layers contained within the cores, Cs^{137} analysis was done on 4 cm slices of sediment. Offshore Sediment Core 3 has a well-defined post 1952 signature in the top 6 cm of the core (Figure 7). A post 1952 signature was not apparent in either of the nearshore cores. In fact, Cs^{137} levels were an order of magnitude less in these cores than in the offshore core and showed no stratification at all. Either the sand was less effective at retaining Cs^{137} or the sediment was disturbed through bioturbation or depositional mixing. Notably, all three cores contained a higher spike of Cs^{137} at their base. I concluded that this was likely due to surface sediment drag down and contamination from the core catcher as it penetrated into the substrate. Despite bottom contamination, offshore sediment core 3 likely preserves an undisturbed sedimentary record.

Discussion:

Quagga mussel and sediment core metaviromes

Quagga mussels are today the dominant invertebrate species in the central offshore Lake Michigan benthos (Pilcher et al., 2017) and this is the first study to document a specific portion of the nanobiome (CRESS-DNA virus-like elements) associated with these invaders. Using genomiphi whole genome amplification methods in the construction of metaviromes produced a known bias towards the discovery of CRESS-DNA virus-like sequences and the consistency in this method across these samples allowing for a comparison of CRESS-DNA viral constituents (Kim and Bae, 2011; Roux et al., 2016). My samples were chosen to produce a

spatially and temporally different look at the CRESS-DNA virus-like sequences. Quagga mussels of different shell length sizes represent different ages with longer and larger quagga mussels being the oldest (Nalepa and Schloesser, 2014). In sediment cores examined in this study the chronosequence begins at the top (youngest) and proceeds to the bottom (oldest). I expected that the viral constituency of all my samples would be quite different due to these spatial and temporal differences as well as the differences in the sample type, a mixture of quagga mussel tissue (comprised of gut, gonads, muscle, and not including the shell) versus sediment comprised of organic detritus originating from both the benthos and falling from the surface waters. Results showed these differences in the overlap in CRESS-DNA viral reads across libraries with a large difference between the sediment and quagga mussel tissue CRESS-DNA virus-like sequences as expected (Figure 6). Even when factoring in total library size, larger differences in reads of CRESS-DNA virus-like sequences were seen between the sediment core layers (SS2 and SS4; Figure 3) then between the different sized quagga mussels which may be explained by the age differences in the two sediment core layers. Differences in the age of the sediment core data is approximately 65 years (extrapolating from our Cs^{137} dating results; Figure 7) versus the difference in time between the large and small quagga mussels representing less than 5 years in time (Nalepa and Schloesser, 2014). A larger amount of time likely allowed for more change both in terms evolution (i.e. rapid sequences changes) and ecology (i.e. changes in host populations) for the CRESS-DNA viral constituency that was preserved in the benthic sediments. For example, although the rep genes of CRESS-DNA viruses are fairly conserved (Rosario et al., 2012) the rate of mutation for other CRESS-DNA viral genes (e.g. capsid genes) has been estimated to be as high $1.2*10^3$ substitutions $site^{-1}$ yr^{-1} (Duffy and

Holmes, 2008; Firth et al., 2009; Grigoras et al., 2010) although these estimates are based on pathogenic CRESS-DNA viruses in which selective pressures may be higher. For the CRESS-DNA viruses found in the benthos of Lake Michigan the recent research in *Diporeia* CRESS-DNA viruses suggests the viruses are not pathogenic which agrees with other studies in aquatic ecosystems that show these viruses to exist asymptomatically in host tissue (Brajae de Oliveriva, 2015). This suggests that the invertebrate CRESS-DNA viruses found in these samples may be under less selective pressure by host systems which could reduce the rate of sequence change.

Focusing on the specific CRESS-DNA virus-like sequences analyzed, reads of nine were detected in both the sediment core samples and the quagga mussel tissue. This indicates that the quagga mussels are interacting with viruses that have been historically present (before 1952 based on Cs^{137} dating results; Figure 7) in the lake that could potentially be associated with other eukaryotes including pelagic and benthic invertebrates, invertebrate associated symbionts, algae, and fungi. BLASTx analysis of these particular virus-like sequences shows them to have similarity with certain algae viral elements, unknown marine viral elements, and sewage viral elements (Table 2). Several of the viruses found in both the quagga mussels and sediment cores were similar (Table 3) to viruses identified in a metavirome study of a sewage treatment oxidation pond (Kraberger et al., 2015). This suggests that these viruses could be allochthonous in origin moving in from the river discharge from the surrounding industrial river mouth sites (Muskegon and Holland) and settling to the benthos were the mussels may interact with them via filter feeding or that they settled in the benthos. Storm and precipitation events combined with lake currents could transport virus-like sequences of terrestrial origin to the

offshore depositional zones in Lake Michigan. These results corroborate with the results of Hewson et al. (2012) who showed the active transport of allochonous viruses into lake ecosystems were they could be detected in sediment samples. However, because none of the CRESS-DNA virus-like sequences identified in this study were complete matches to these sequences and due to the rapid evolutionary rates of CRESS-DNA viruses and all the unknown sequence information still out there it is difficult to speculate about the potential ecology of these CRESS-DNA virus-like sequences identified in the metavirome libraries (Rosario et al. 2012). It should also be noted that bioturbation may play a role in moving and distributing viruses in the sediments core layers although the Cs^{137} dating results of the offshore sediment core 3 shows a distinct Cs^{137} peak and then low levels throughout the rest of the core suggesting minimal disturbance.

LMQMvls1241 qPCR

In addressing my hypothesis regarding what CRESS-DNA virus-like sequences quagga mussels may have moved into the lakes when they established significant populations, my data does not strongly support the notion that specific viruses were brought in and established in high abundance. The qPCR analysis of LMQMvls1241 showed it to be prevalent in the offshore benthos of Lake Michigan but below the detection limit in the inland waters of Muskegon Lake or in the Rybinsk Reservoir in Russia. LMQMvls1241 was chosen for further analysis because it was a well-represented CRESS-DNA virus Rep gene (most similar to the rep gene of a Bat circovirus) with several ORFs to target with qPCR and it was not present in the sediment cores so the origins were unknown. Several hypotheses could explain why LMQMSvls1241 was detected only in the offshore benthos of Lake Michigan. First, the CRESS-DNA may have

infected another eukaryotic species that was present in high abundance in the offshore benthos of Lake Michigan or is part of the filter feeding diet of the mussels. A very recent study by Bistolas et al. (2017c) investigated specific CRESS-DNA viruses found in marine isopods across the west coast of the United States and was able to show that when particular ciliate epibionts (unicellular eukaryotes) were mechanically removed from the isopods and specific CRESS-DNA virus they were targeting with qPCR was no longer detected indicating that ciliate epibionts are likely the putative host of the CRESS-DNA virus. In dreissenid mussels a particular commensal ciliate, *Conchopthirus acuminatus,* is highly prevalent in mussels in the native range but has not been found in dreissenid mussels in North America (Karatayev et al., 2000; Conn et al., 2008). No research has recently looked at the ciliate endosymbionts in the recent large populations of Great Lakes quagga mussels but research looking into these endosymbionts and potential associated CRESS-DNA viruses could help to understand more about invertebrate associated CRESS-DNA viral ecology. The second hypothesis is that LMQMSvls1241 is specific to the profunda morph genetic strain of the quagga mussel which is adapted to life in the deep offshore waters of the Great Lakes (Dermott and Munawar, 1993; Claxton et al., 1998; Nalepa et al., 2009). This is supported by the study of Bistolas et al. (2017) who showed *Diporeia* associated CRESS-DNA viral abundance to be associated with specific haplotype demographics of *Diporeia spp.* with the highest abundance found in the specific southern clade. The third hypothesis is that LMQMvls1241 was a rare virus present in quagga mussel tissues or in the diet of the quagga mussels and conditions in the offshore benthos with high population densities and growth allowed for the virus to become abundant. Although LMQMvls1241 was below any detection limits in Muskegon Lake and Russian samples in the qPCR reactions, there were

inhibition impacts and as a consequence it cannot be ruled out that the virus-like sequence could be present in low abundance in the samples from Muskegon Lake and the Rybinsk Reservoir in Russia and as result, characterizing the origins of LMQMvls1241 remains to be discovered. Finally it should be noted that I was comparing a small sample of quagga mussels and focusing on only present day quagga mussels and a small spatial area where quagga mussel populations are found in both North America and Europe. Sampling of a greater number of quagga mussels over a much wider spatial and temporal range is necessary to make any predictions about the origins and distribution of particular CRESS-DNA virus-like sequences.

In comparison of the measured viral load of *Diporeia* associated LM29173 circoviruses, the dominate virus in metavirome libraries and targeted qPCR in Hewson et al. (2013) and Bistolas et al. (2017), our values for LMQMvls1241 were much lower than those measured in Lake Michigan *Diporeia* but closer to the average value measured in Lake Superior *Diporeia* (Table 4). Compared to less common *Diporeia* associated CRESS-DNA viruses (LM122 and LH481) in these studies (Hewson et al., 2013 and Bistolas et al., 2017) metavirome libraries and then quantified using qPCR, LMQMvls1241 viral load was higher on average (Figure 8). These comparisons should be taken cautiously as the sample size in my study was considerably smaller. As far as prevalence, LMQMvls1241 was detected in all quagga mussel samples collected from the offshore benthos of Lake Michigan. These results suggest that LMQMvls1241 is a common constituent of the quagga mussel nanobiome in the offshore benthos of Lake Michigan but likely not the dominate CRESS-DNA virus present in the quagga mussels.

Another possibility remains that other CRESS-DNA virus-like sequences that were not observed via qPCR were transported in the quagga mussels when they invaded. There were

many CRESS-DNA virus-like sequences detected in the quagga mussel tissues that were not found preserved in the two different sediment layers. Future work should include a metaviromic analysis of CRESS-DNA virus-like sequences in quagga mussel samples from the native range in the Ponto-Caspian Sea region as well as those first samples of quagga mussels that established in the Great Lakes to offer a comparison of the viruses that are shared and could have been moved into the Great Lakes. Ideal sampling areas may include international shipping ports that would likely be the transfer points for the invasive mussels.

This study provides a broad look at the temporal and spatial variation in CRESS-DNA viruses in the benthos of large freshwater lake helping to address a few key aspects of the ecology of these viruses in freshwater systems. A review by Middleboe et al. (2008) emphasized the need for attention on benthic viral ecology and this study and others looking at benthic invertebrates in the Great Lakes (Hewson et al., 2013; Bistolas et al., 2017a; 2017b) has helped describe what is happening to freshwater ecosystems with the presence of such viruses. This study shows that CRESS-DNA virus-like genotypes found preserved in sediments are still present today and circulating in current populations of invasive species. In the Great Lakes, evidence shows that CRESS-DNA viruses appear to not be pathogenic or impact invertebrate host (*Diporeia spp.*) physiology, yet they are common constituents in the host nanobiome (Bistolas et al., 2017a; 2017b). A wide spread survey of CRESS-DNA viruses in a New Zealand lake system revealed that the CRESS-DNA viruses identified and detected were widespread in lake ecosystems and not simply associated with invertebrate hosts analyzed (Dayaram et al., 2016). In studies by Dayaram et al. (2013; 2016), molluscs and filter feeding bivalves were found to have the greatest diversity of CRESS-DNA viruses which may be explained by their

filtering feeding behavior which allows them to interact with and potentially concentrate CRESS-DNA virus circulating in a lake ecosystem. It has been hypothesized that many CRESS-DNA viruses are actually infecting algal and fungal eukaryotic hosts which are widespread, extremely common, and diverse constituents in freshwater benthic systems (very common in sediments) where fungi are important decomposers of organic matter (Krauss et al., 2011; Dayaram et al., 2016). Future studies interested in the ecology of these viruses should focus on algal, fungal and other eukaryotic symbionts that invertebrates in freshwater systems may interact with and may be the true hosts of many of these invertebrate associated CRESS-DNA viruses.

Conclusions

In this study I hypothesized that invasive species like quagga mussels would have hosted novel CRESS-DNA viruses that could have potential interesting ecological consequences. This study, however, shows that CRESS-DNA virus-like genotypes historically present in Lake Michigan benthic sediments are interacting with the invasive quagga mussels in the Lake Michigan benthos indicating the wide distribution of particular native CRESS-DNA virus-like genotypes both ecologically, temporally, and spatially. Although the difference between the CRESS-DNA viral consortium found in the sediment core layers and in the quagga mussels tissue was large, there is not definitive evidence to suggest that quagga mussels hosted and transported any new virus-like genotypes. This study does show that LMQMvls1241 was likely not brought in with the invasive quagga based on its absence in Russian quagga mussel samples. As a consequence, this study offers new insight into how CRESS-DNA viruses are preserved in the benthos, move through the food webs, and how invasive species are

interacting with these viruses. Future research should complete metavirome libraries of different Great Lakes invaders to those in the native range to offer a true comparison of the similarity of CRESS-DNA viruses found in the same organisms moving between ecosystems.

References:

Barbiero, R.P., Schmude, K., Lesht, B.M. Riseng, C.M., Warren, G.J., Tuchman, M.L., 2011. Trends in *Diporeia* populations across the Laurentian Great Lakes, 1997-2009. Journal of Great Lakes Research 37, 9-17.

Bistolas, K.S.I., Besemer, R.M., Rudstam, L.G., Hewson, I., 2017c. Distribution and Inferred Evolutionary Characteristics of a Chimeric ssDNA Virus Associated with Intertidal Marine Isopods. Viruses 9, 361.

Bistolas, K.S.I., Jackson, E.W., Watkins, J.M., Rudstam, L.G., Hewson, I., 2017a. Distribution of circular single-stranded DNA viruses associated with benthic amphipods of genus *Diporeia* in the Laurentian Great Lakes. Freshwater Biology 62(7), 1220-1231.

Bistolas, K.S.I., Rudstam, L.G., Hewson, I., 2017b. Gene expression of benthic amphipods (genus: *Diporeia*) in relation to a circular ssDNA virus across two Laurentian Great Lakes. PeerJ, e3810; DOI 10.7717/peerj.3810.

Brajao de Oliveria, K. 2015. Torque teno virus; a ubiquitous virus. Revista Brasileira de Hematolgia e Hemterpia 37(6), 357-358.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology 13, 278-284.

Brown, J., 2014. Comparison of water column virus populations over a 180+ year time period within the varved sediments of Fayetteville Green Lake. Chapter 4 of Dissertation. Cornell University.

Cave, C.S., Strychar, K. B. 2014. Decline of *Diporeia* in Lake Michigan: Was Disease Associated with Invasive Species the Primary Factor? International Journal of Biology 7, 93.

Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18, 117-143.

Claxton, W.T., Wilson, A.B., Mackie, G.L., Boulding, E.G., 1998. A genetic and morphological comparison of shallow and deepwater populations of the introduced dreissenid bivalve *Dreissena bugensis* Canadian Journal of Zoology 76, 1269-1276.

Cook, D.G., Johnson, G., 1974. Benthic macroinvertebrates of the St Lawrence Great Lakes. Journal of Fisheries Research Board of Canada 31, 763-782.

Coolen, M.J., 2011. 7000 years of Emiliania huxlyi viruses in the Black Sea. Science 33, 451-452.

Coon, D.B., Simpson, S.E., Minchin, D., Lucy, F.E., 2008. Occurrence of *Conchophthirus acuminatus* (Protista: Ciliophora) in *Dreissena polymorpha* (Mollusca: Bivalvia) along the River Shannon, Ireland. Biological Invasions 10, 149-156.

Coventry, R. J., Fett, D.E.R., 1979. A Pipette and Sieve Method of Particle-size Analysis and Some Observations on Its Efficacy. CSIRO Report No.38, 26p.

Cuhel, R.L., Aguilar, C., 2013. Ecosystem Transformations of the Laurentian Great Lake Michigan by Nonindigenous Biological Invaders. Annuals Review in Marine Sciences 5, 289-320.

Dayaram, A., Goldstein, S., Zawar-Reza, P., Gomez, C., Harding, J.S., Varsani, A., 2013. Identification of starling circovirus in an estuarine mollusc (Amphibola crenata) in New Zealand using metagenomics approaches. Genome Announcement 1(3), e0027-13.

Dayaram, A, Goldstien, S, Arguello-Astroga, GR, Zawar-Reza, P., Gomez C., Harding J.S., Varsani, A. 2015. Diverse small circular DNA viruses circulating amongst estuarine molluscs. Infections, Genetics and Evolution 31, 284-295.

Dayaram, A., Galtowitsch, M.L., Arguello-Astorga, G.R., Van Bystervelt, K., Kraberger, S. Stainton, D., Harding, J.S., Roumagnac, P., Martin, D.P., Lefeuvre P., Varsani A., 2016. Diverse circular replication-associated protein encoding viruses circulating within a lake ecosystem. Infection, Genetics and Evolution 39, 304-316.

Dermott, R., Munawar, M., 1993. Invasion of Lake Erie offshore sediments by Dreissena, and its ecological implications. Canadian Journal of Fisheries and Aquatic Sciences. 50, 2298-2304.

Dermott, R., Munawar, M., Bonnell, R., Carrou, S., Nibblock, H., Nalepa, T., Messick, G., 2005. Preliminary investigations for the causes of the disappearance of *Diporeia spp*. from Lake Ontario. Pp. 203-232 in Great Lakes Fisheries Commission technical report. Great Lakes Fisheries Commission. Ann Arbor, MI.

Dunlap, D. S., Ng, T. F. F., Rosario, K., Barbosa, J. G., Greco, A. M., Breitbart, M., Hewson, I., 2013. Molecular and microscopic evidence of viruses in marine copepods. Proceedings of the National Academy of Sciences, 110(4), 1375-1380.

Duffy, S., Holmes, E.C., 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begmoviruses tomato yellow leaf curl virus. Journal of Virology 82, 957-965.

Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. National Review of Microbiology 3, 504-510.

Elston, R., 1997. Special topic review: Bivalve mollusk viruses. World Journal of Microbiology and Biotechnology 13, 393-403.

Firth, C., Chartleston, M.A., Duffy S., Shapiro B., Holmes E.C., 2009. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. Journal of Virology 83, 12813-12821.

Fitzgerald, S.A., Gardner, W.S., 1993. An algal carbon budget for pelagic-benthic coupling in Lake Michigan. Limnology and Oceanography 38(3), 547-560.

Fuhrman, J. A., 1999. Marine viruses and their biogeochemical and ecological effects. Nature, 399(6736), 541-548.

Gastrich, M.D., Leigh-Bell, J.A., Gobler, C.J., Anderson, O.R., Wilhelm, S.W., Bryan, M., 2004. Viruses as potential regulators of regional brown tide blooms caused by the alga, *Aureococcus anophagefferens.* Estuaries 27(1), 112-119.

Grigoras, I., Timchenko T., Grande-Perez, A., Katul, L., Vetten, H.J., Gronenborn, B., 2010. High variability and rapid evolution in nanovirus. Journal of Virology 84, 9105-9117.

Haywick, D.W., 2006. Manual of Pipette and Sieve Particle Analysis. University of South Alabama, 35pp.

Hewson, I., Barbosa, J.G., Brown J.M., Donelan, R.P., Eaglesham, J.B., Eggleston, E.M., LaBarre, B.A., 2012. Temporal dynamics and decay of putatively allochthonous and autochthonous viral genotypes in contrasting freshwater lakes. Applied and Environmental Microbiology 78(18), 6583-6591.

Hewson, I. Eaglesham, J.B., Hook, T.O. LaBarre, B.A., Sepulveda, M.A., Thompson, P.D., Watkins, J.M., Rudstam, L.G., 2013a. Investigation of viruses in *Diporeia spp*. from the Laurentian Great Lakes and Owasco Lake as potential stressors of declining populations. Journal of Great Lakes Research 39, 499-506.

Hewson, I., Fuhrman, J.A., 2003. Viriobenthos production and virioplankton sorptive scavenging by suspended particles in coastal and pelagic waters. Microbiology Ecology 46, 337-347.

Hewson, I. Ng, G., Li, W., LaBarre, B.A., Aguirre, I., Barbosa, J.G., Breitbart, M., Greco, A.W., Kearns, C.M., Looi, Al, Schaffner, L.R., Thompson, P.D., Hairston, N.G., 2013b. Metagenomic identification, season dynamics and potential transmission mechanisms of a Daphnia- associated single- stranded DNA virus in two temperate lakes. Limnology and Oceanography 58(5), 1605-1620.

Hurst, C.J., 2000. Viral Ecology. Academic Press. 659p.

Karatayev, A.Y., Molloy, D.P., Burlakova, L., E., 2000. Seasonal Dynamics of *Conchophtirus acuminatus* (Ciliophora, Conchophthirdiae) Infection in *Dreissena polymorpha* and D. bugensus (Bivalvia, Dreissenidae). European Journal of Protistology 36, 397-404.

Karatayev, A.Y., Burlakova L.E., Padilla D.K., 2015. Zebra versus quagga mussels: a review of their spread, population dynamics, and ecosystems impacts. Hydrobiologia 746, 97-112.

Keller, R. P., Cadotte, M. W., Sandiford, G., 2014. Invasive Species in a Globalized World. University of Chicago Press. 427p.

Kim, K.H., Bae, J.W., 2011. Amplification methods bias metagenomic Libraries of uncultured single-stranded and double-stranded DNA viruses. Applied and Environmental Microbiology 77, 7663-7668.

Kraberger, S., Arguello-Astorga, G.R., Greenfield, L.G., Galilee, C., Law, D., Martin, D.P., Varsani, A., 2015. Characterization of a diverse range of circular replication associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. Infection, Genetics, Evolution. 31, 73-86.

Krauss, G.J., Sole, M., Krauss, G., Schlosser, D. Wesenberg, D. Barlocher F., 2011. Fungi in freshwaters: ecology, physiology and biochemical potential. FEMS Microbiology Review 35, 620-651.

Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., Alcami, A., 2009. High Diversity of the Viral Community from an Antarctic Lake. Science News Series 326, 858-861.

Lorenz, M.G., Aardema, B.W., Krumbein, W.E., 1981. Interaction of marine sediments with DNA and DNA availability to nucleases. Marine Biology 64, 225-230.

Lymer, D., Logue J.B., Brussard, C.P.D., Baudoux, A.C., Vrede K., Lindstrom, E.S., 2008. Temporal variation in freshwater viral and bacterial community composition. Freshwater Biology 53, 1163-1175.

Marzolf, R. G., 1965. Substrate Relations of the Burrowing Amphipod Pontoporeia Affinis in Lake Michigan. Ecology 46(5), 580-592.

Middleboe, M., Jacquet, S., Weinbauer, M., 2008. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. Freshwater Biology 53, 1069-1075.

Murray, A.G., Jackson, G.A., 1992. Viral dynamics: a model of the effects size, shape, motion and abundance of single-celled planktonic organisms and other particles. Marine Ecology Program Series 89, 103-116.

Nagasaki, K., Tarutani, K., Tomaru, Y., Itakura, S., Tamai, K., Yamanaka, S. Yamaguchi, M., 2002. Possible use of viruses as a microbiological agent against harmful algal blooms. Fisheries Science 68, 497-500.

Nalepa, T.F., Hartson, D.J., Fanslow, D.L., Lang, G.A., Lozano, S.J., 1998. Declines in benthic macroinvertebrate populations in southern Lake Michigan, 1980-1993. Canadian Journal of Fisheries and Aquatic Science 55, 2402-2413.

Nalepa, T.F., Fanslow, D.L, Iii, A.J.F., Lang, G.A., Eadie, B.J., Quigley, M.A., 2006. Continued disappearance of the benthic amphipod *Diporeia spp*. in Lake Michigan: is there evidence for food limitation? Canadian Journal of Fisheries and Aquatic Science 63, 872-890.

Nalepa, T.F., Fanslow, D.L., Lang, G.A., 2009. Transformation of the offshore benthic community in Lake Michigan: a recent shift from the native amphipod *Diporeia spp.* to the invasive mussel *Dreissena rostriformis bugensis*. Freshwater Biology 54, 466-479.

Nalepa, T.F., Schloesser, D.W., 1993. Zebra Mussels Biology, Impacts and Control. Lewis Publishers, Boca Raton, 810p.

Nalepa, T.F., Schloesser, D.W., 2014. Quagga and Zebra Mussels: Biology, Impacts and Control. 2nd edition. CRC Press, Boca Raton, 7750p.

Nalepa, T.F., Fanslow, D.L., Lang G.A., Mabrey, K., Rowe M., 2014. Lake-wide benthic surveys in Lake Michigan 1994-95, 2000, 2005, and 2010: Abundances of the amphipod *Diporeia spp.* and abundances and biomass of the mussels *Dreiessena polymorpha* and *Dreissena rostiformis bugensis* 21 pp. NOAA Technical Memorandum GLERL -164, NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, Mich.

Noble, R.T., Fuhrman, J.A., 1998. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquatic Microbial Ecology 14, 113-118.

Ng, T.F.F., Willner, D., Nilsson, C., Lim, Y.W., Schmieder, R., Chau, B., Ruan, Y., Rohwer, F., Breitbart, M., 2010. Vector-based metagenomics for animal virus surveillance. International Journal Infectious Disease 14, E378.

Pagnucco, K.S., Maynard, G.A., Fera, S.A., Yan, N.D., Nalepa T.F., Ricciardi A., 2015. The future of species invasions in the Great Lakes-St. Lawrence River basin. Journal of Great Lakes Research 41, 96-107.

Pilcher D.J., Mckinley G.A., Kralj J., Bootsma H.A., Reavie E.D., 2017. Modeled sensitivity of Lake Michigan productivity and zooplankton to changing nutrient concentrations and quagga mussels. Journal of Geophysical Research: Biogeosciences 122, 2017-2032.

Rosario, K., Duffy, S., Breitbart, M., 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. Journal of General Virology 90, 2418-2424.

Rosario, K. Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic single-stranded DNA viruses: insights gained from metagenomics. Archives of Virology 157, 1851-1871.

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T., Debroas, D., 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. PLoS One 7, e33641.

Roux, S., Solonenko, N.E., Dang, V.T., Poulos, B.T., Schwenck, S. M., Goldsmith, D.B., Sullivan, M.B., 2016. Towards quantitative viromics for both double-stranded and singles stranded DNA viruses. PeerJ, 4, e2777.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and biologists programmers. In: Krawertz. S., Misener S. (Eds.), Bioinformatics Methods and Protocols, Humana Press., Totowa,NJ, pp. 365-386.

Sly, P.G., W.J. Christie., 1992. Factors influencing densities and distributions of Pontoporeia hoyi in Lake Ontario. Hydrobiologia 235, 321-352.

Sanctis, D.M., Adeeb A., Farcomeni A. Patriarca C., Saed A. & Attore F., 2013. Classification and distribution patterns of plant communities on Socotra Island, Yeman. Applied Vegetation Science 16, 148-165.

Short, C. M., Oksana R., Short S.M., 2011. Quantification of virus genes provides evidence for seed-bank populations of phycodnaviruses in Lake Ontario, Canada. International Society for Microbial Ecology 5, 810-821.

Short, S.M. 2012. The ecology of viruses that infect eukaryotic algae. Environmental Microbiology 14(9), 2253-2271.

Suttle, C.A., 2007. Marine viruses- major players in the global ecosystem. National Review of Microbiology 5, 801-812.

Suttle, C.A., 2013. Viruses: unlocking the greatest biodiversity on Earth. Genome 10, 542-544.

Thingstad, T.F., 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnology and Oceanography 45, 1320-1328.

Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. Nature protocols 4(4), 470-483.

Van Etten, J.L., Van Etten, C.H., Johnson, J.K., Burbank D.E., 1985. A survey for viruses from freshwater that infect a eukaryotic Chlorella-like green alga. Applied Environment Microbiology 49, 1326-1328.

Vanderploeg, H.A, Nalepa, T.F., Jude, D.J., Mills, E.L., Holeck, K.T., Liebig, J.R., Grigorovich, I.A., Ojaveer, H., 2002. Dispersal and emerging ecological impacts of Ponto- Caspian species in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 59, 1209-1228.

Vanderploeg, H.A. Liebid, J.R., Nalepa, T.F., Fahnenstiel, G.L., Pothoven, S.A., 2010. Dreissena and the disappearance of the spring phytoplankton bloom in Lake Michigan. Journal of Great Lakes Research 36, 50-59.

Watkins, J.M., Dermott, R., Lozano, S.J., Mills, E.L. Rudstam, L.G., Scharold, J.V., 2007. Evidence for remote effects of dreissenid mussels on the amphipod *Diporeia*: analysis of Lake Ontario Benthic Surveys, 1972-2003. Journal of Great Lakes Research 33, 642-657.

Watkins, J.M., Rudstam, L.G., Mills, E.L., and Teece, M.A., 2012. Coexistence of the native benthic amphipod *Diporeia spp.* and exotic dreissenid mussels in the New York Finger Lakes. Journal of Great Lakes Research 38, 226-235.

Wilhlem, S.W., Suttle, C.A. 1999. Viruses and nutrient cycles in the sea. Bioscience 49, 781-788.

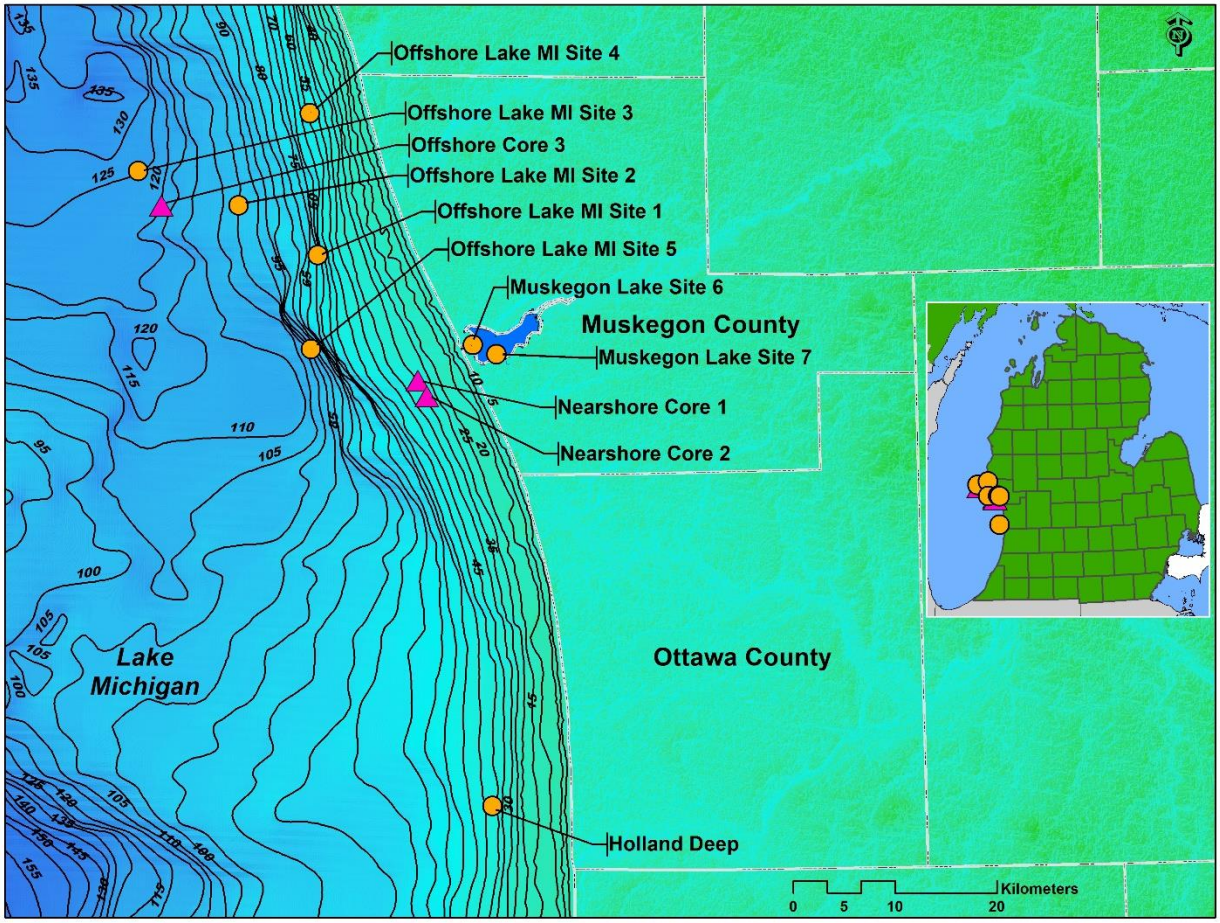
Winnell, M.H., White, D.S., 1984. Ecology of Shallow and Deep Water Populations of Pontoporiea hoyi (Smith) (Amphipoda) in Lake Michigan. Freshwater Invertebrate Ecology 3, 118-138.

Winters, A.D., Fitzgerald, S., Brendan, T.O., Napela, T., Faisal, M., 2014. Spatio-temporal dynamics of parasites infecting *Diporeia spp.* (Amphipoda, Gammaridae) in southern Lake Michigan (USA). Journal of Invertebrate Pathology 121, 37-35.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., Takai, K., 2013. Metagenomic Analysis of Viral Communites in (Hado) Pelagic Seidments. PLoS One 8(2), e5727.

Figures and Tables:

(A)



(B)

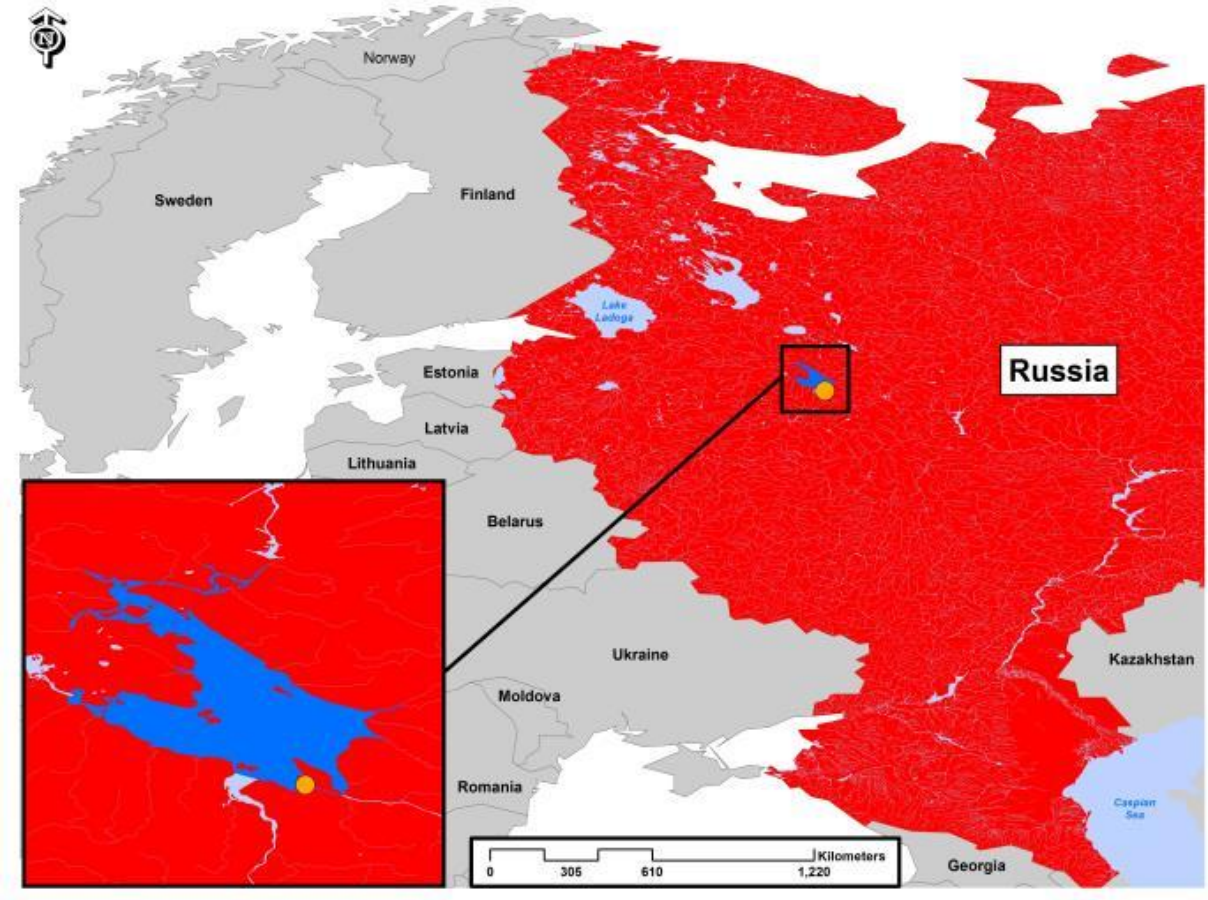


Figure 1: The top figure (A) shows a bathymetric map (NOAA) of the study area in the central Lake Michigan benthos with the quagga mussel and sediment core sampling locations labeled. Number values along contour lines represent the water depth in meters. Pink triangles indicate the locations of sediment core sampling in the summer of 2016. Nearshore Core 1 and Core 2 represents the location of nearshore sediment cores collected with the vibracore on July 19, 2016. Offshore Core 3 represents the location of the offshore gravity core collected on July 21, 2016. Orange circles indicate the locations where quagga mussels were collected in ponar grabs in the summer and fall of 2016. The Holland Deep location represents the initial quagga mussel sampling location used for the metavirome collected on June 23, 2107. The Lake Michigan sites 1-5 and Muskegon Lake sites 6-7 represent quagga mussel sampling for targeted qPCR analysis collected on September 24, 2016. The bottom figure (B) shows a map of the location of the Rybinsk Reservoir along the Volga River in Russia where quagga mussel samples were collected in July 2017. The orange circle indicates the location where the quagga mussels were collected.

Figure 2: Picture of sediment core subsampling showing the division of the core into eight sections labeled SS1 (top of the core) to SS8 (bottom of the core). The red arrow from SS4 shows the location where sediment was sampled for viral analysis.

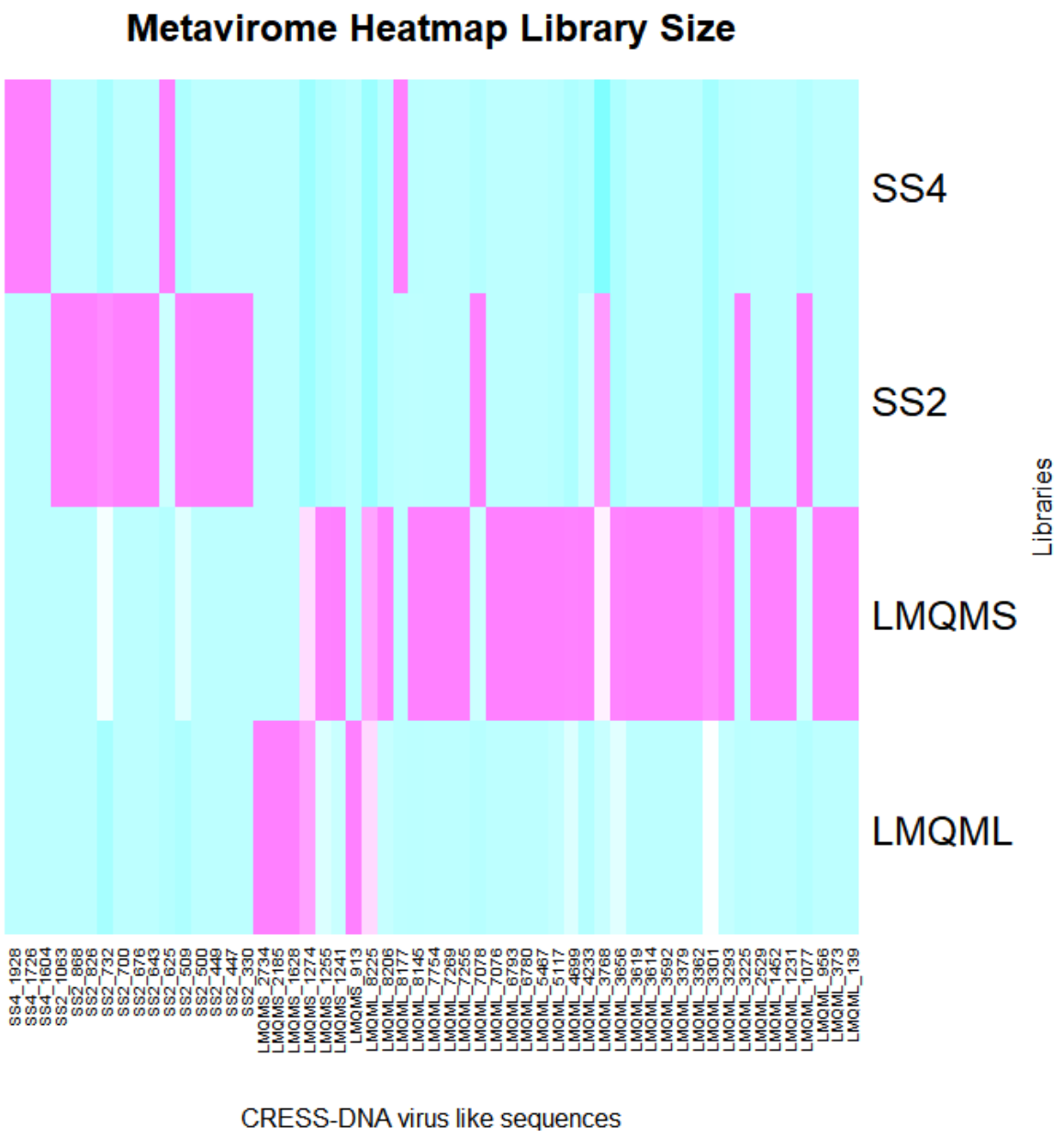


Figure 3: A heat map representing the number of reads of the 54-CRESS-DNA virus-like sequences found across the four metavirome libraries created in this study relative to library size. LMQML represents the large quagga mussel library, LMQMS represents the small quagga mussel library, SS2 represents the upper layer (younger) in offshore sediment core 3, and SS4 represents the lower layer (older) in offshore sediment core 3. The CRESS-DNA virus-like sequences are labeled with the library in which they were first identified and then the contig number. To factor in library size the number of reads detected of each CRESS-DNA virus-like sequence was divided by the total library size. Light blue represents a low number of reads where pink and purple represents a high number of reads detected.

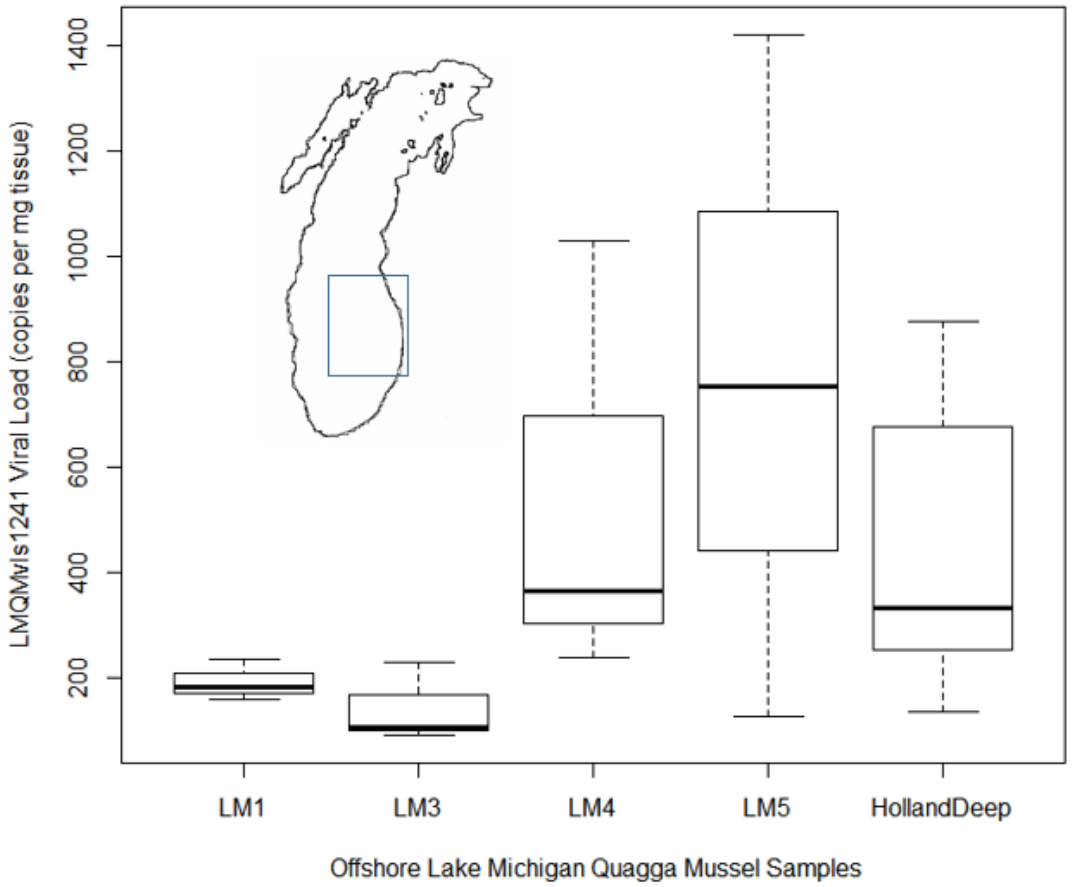


Figure 4: Lake Michigan Quagga Mussel CRESS-DNA virus-like sequence contig 1241 (LMQMvls1241) qPCR analysis of quagga mussel samples from offshore Lake Michigan. The vertical axis represents LMQMvls1241 viral load (copies of LMQMLvls1241 rep ORF per mg of quagga mussel tissue) and the horizontal represents all the locations of samples. LM1-5 represents quagga mussels collected in fall 2016 from offshore Lake Michigan sites 1 – 5 (Site 2 was not included because of low sample size n=1; Figure 1A). Three individual quagga mussels (of variable size) were analyzed from each site (n=3). Holland Deep represents quagga mussels collected in summer 2016 from the Holland Deep location (Figure 1A). Six individual quagga mussels, two from each size class (large >25mm SL, medium 15mm-25mm SL, and small <15mm SL) were analyzed from this site.

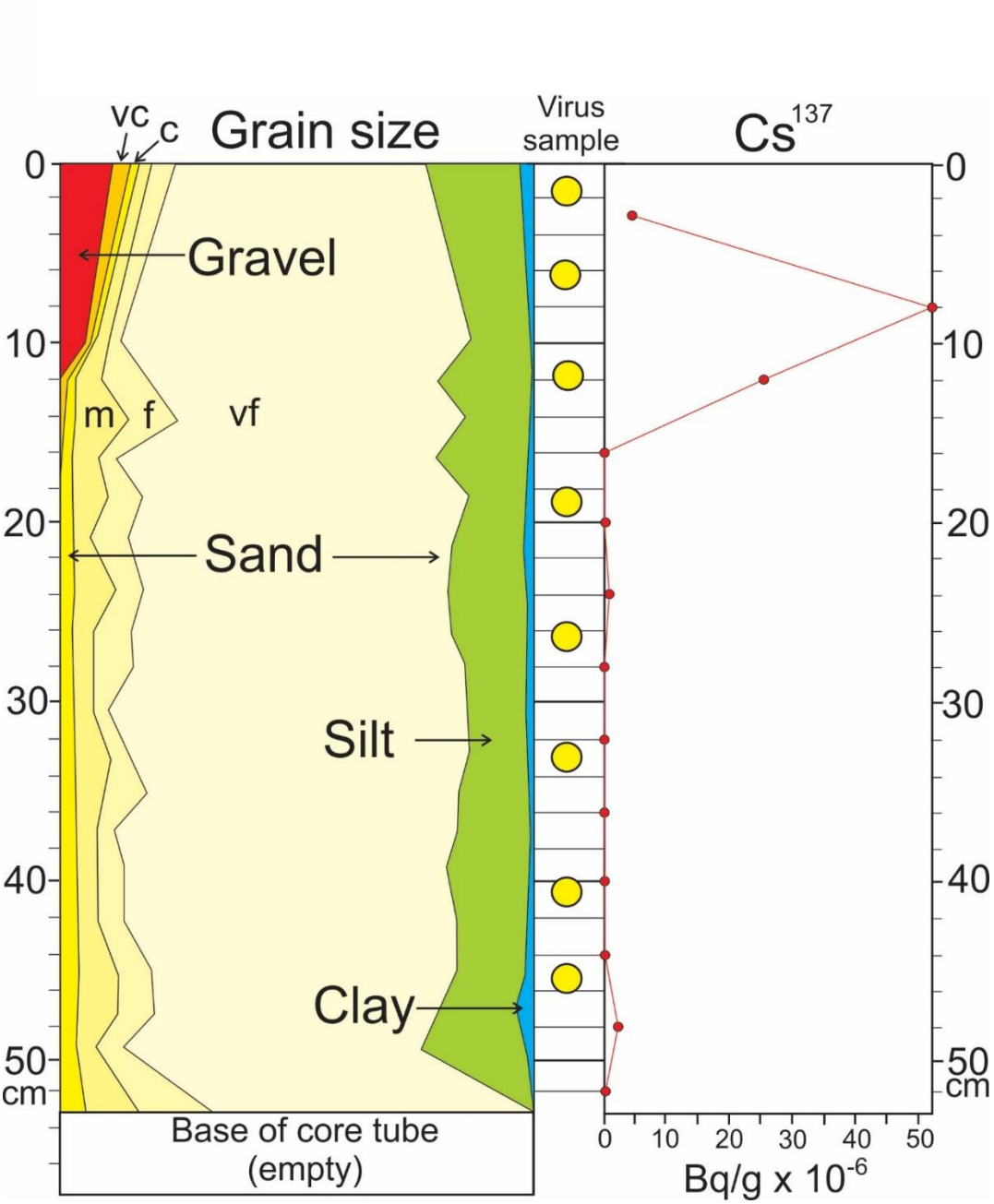


Figure 5: Diagram showing the grain size analysis from Lake Michigan nearshore sediment core site 1 core 1(Figure 1) fractions starting at the top of the core (0 cm = the bottom of the lake) and moving down into the earth (50cm +) . Red= gravel substrate, yellow = sand substrate, green = silt substrate, and blue corresponds = clay substrate. The sand substrate is shown in higher resolution where VC= very course, C= coarse, F=fine, VF= very fine. The right side of the figure represents the Cs^{137} dating analysis that measured the fraction of Cs^{137} in the sediment layers. The dotted line represents the year 1952 when atmospheric Cs^{137} levels were at a maximum from nuclear weapons testing. The gray at the bottom represents contamination in the core from the sampling process.

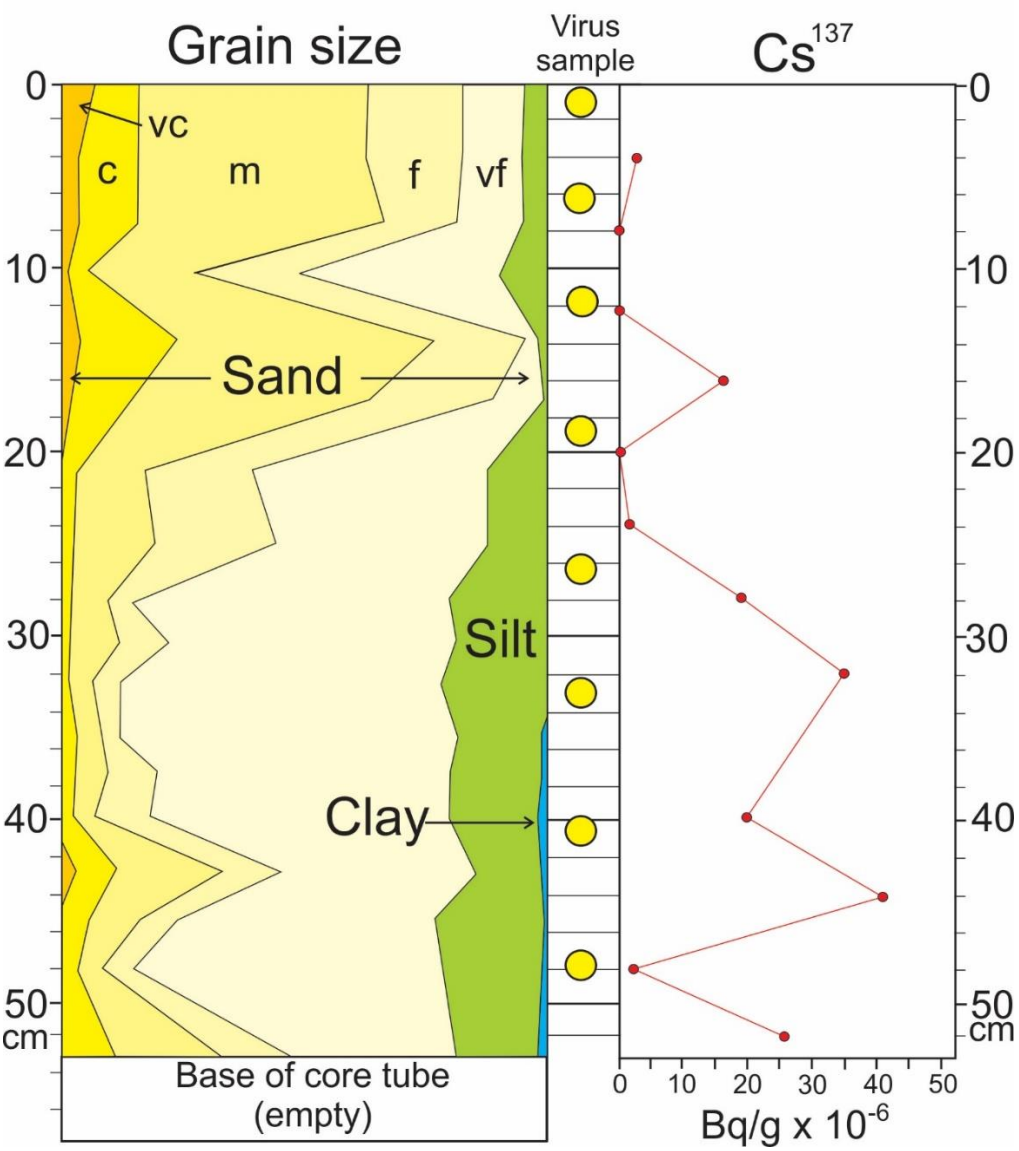


Figure 6: Diagram showing the grain size analysis from Lake Michigan nearshore sediment core site 1 core 2 (Figure 1) fractions starting at the top of the core (0 cm = the bottom of the lake) and moving down into the earth (50cm +) . Red= gravel substrate, yellow = sand substrate, green = silt substrate, and blue corresponds = clay substrate. The sand substrate is shown in higher resolution where VC= very course, C= coarse, F=fine, VF= very fine. The right side of the figure represents the Cs^{137} dating analysis that measured the fraction of Cs^{137} in the sediment layers. The dotted line represents the year 1952 when atmospheric Cs^{137} levels were at a maximum from nuclear weapons testing. The gray at the bottom represents contamination in the core from the sampling process.

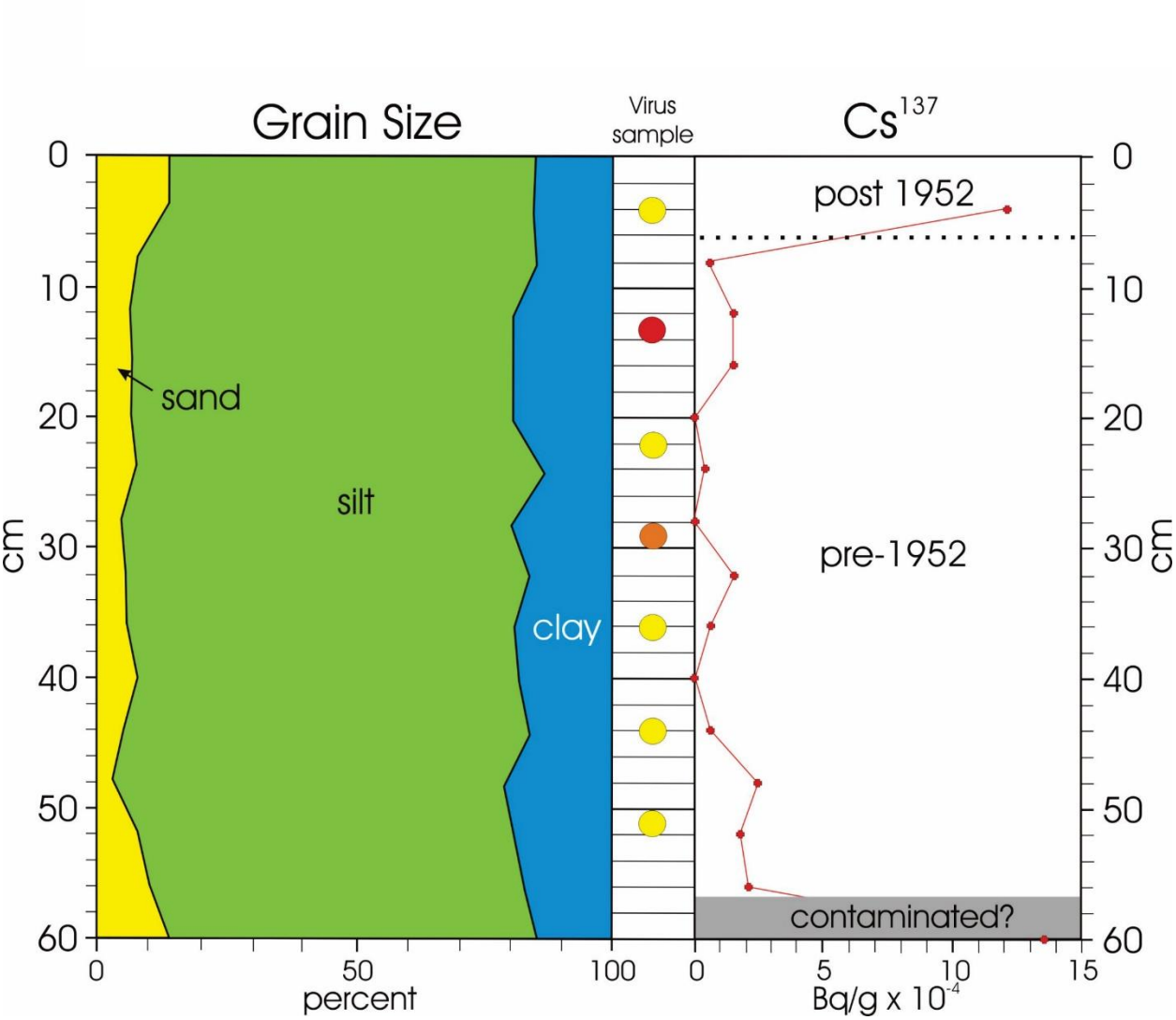


Figure 7: Left side represents the grain size analysis of offshore sediment core 3(Figure 1) fractions starting from the top of the core (0cm= bottom of the lake) and going deeper into the earth (50cm +). The yellow color = sand substrate, the green color = silt substrate, and the blue color = clay substrate. The middle figure that is labeled virus sample represents all the sediment core subsample locations (SS1-SS8, with SS1 being at the top and SS8 at the bottom) that were sampled for analysis both qPCR (LM29173) and metaviromics. The red spot represents subsample SS2 and the orange spot represents subsample SS4 in the metaviromic analysis. The right side of the figure represents the Cs^{137} dating analysis that measured the fraction of Cs^{137} in the sediment layers. The dotted line represents the year 1952 when atmospheric Cs^{137} levels were at a maximum from nuclear weapons testing. The gray at the bottom represents contamination in the core from the sampling process.

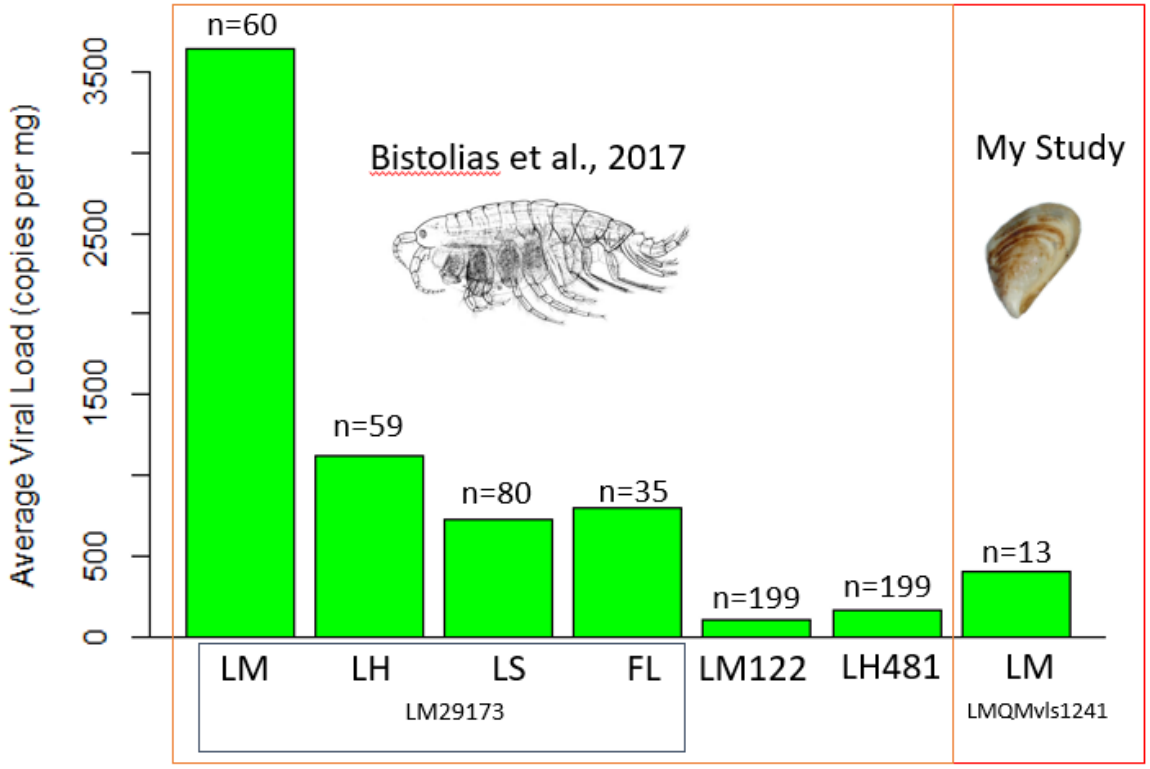


Figure 8: The average viral load for CRESS-DNA viral genotypes identified in studies in the Great Lakes and Finger Lakes (Bistolas et al., 2017; *Diporeia spp.*) including my qPCR results for LMQMvls1241. The average viral load of *Diporeia* associated LM29173 circoviruses is shown in the left four bars where LM stands for Lake Michigan, LH stands for Lake Huron, LS stands for Lake Superior, and FL stands for Finger Lakes. The results from the other *Diporeia* associated circoviruses, LM122 and LH481 are given in the fifth and sixth bars respectively with samples pooled from across the Great Lakes region. Finally my results for LMQMvls1241 in Lake Michigan quagga mussels is shown in the red box on the far right. The sample size (number of organisms in which qPCR was run) is above each bar.

Table 1: Sample descriptions, date of collection, location of collection, the depth of water (m) collected in, the size of the metavirome libraries, total contigs, and CRESS-DNA virus-like sequences for quagga mussel and sediment core subsample investigated in the metaviromes completed in this study. In the sample descriptions, five mussels were used for each of the quagga mussel samples and a recorded mass of sediment taken at a recorded depth down in the core from the surface. SL = shell length which was measured from the end of the shell attachment point to the furthest tip.

Metavirome Library Sample Descriptions	**Date Collected**	**Location (Figure 1)**	**Water Depth(m)**	**Total Library Size after trimming**	**Total Contigs**	**Novel CRESS-DNA virus-like sequences**	**CRESS-DNA virus-like sequence reads**
LMQML: Large Quagga Mussels (5 mussels with > 25 mm SL)	6/23/2016	Holland Deep	47	1705104	3076	31	52
LMQMS: Small Quagga Mussels (5 mussels with < 15 mm SL)	6/23/2016	Holland Deep	47	3071512	8417	7	23
SS2: Sediment Core Subsample 2 (1.27g of sediment -136 mm down in core)	7/21/2016	Offshore Sediment Core 3	120	3565172	1081	13	16
SS4: Sediment Core Subsample 4 (4.55g of sediment -290 mm down in core)	7/21/2016	Offshore Sediment Core 3	120	3225506	3629	3	16

Table 2: CRESS-DNA virus-like sequences of interest identified in metavirome libraries BLASTx (NCBI) results summary. The table displays the CRESS-DNA virus-like sequence, the 1st BLASTx hit, 2nd BLASTx hit, and 1st relevant freshwater CRESS-DNA virus BLASTx hit, and information on query cover, e-value, % identity and Accession # for each CRESS-DNA viral element listed. For the 1st relevant freshwater CRESS-DNA virus BLASTx hits the place in the list in which it is found is listed in the far right column.

Virus Like Sequence	**1st Closest BLASTx HIT**	**Query Cover**	**e-value**	**% identity**	**Accession #**	
LMSS2vls625	Rep domain protein [Phaeocystis globosa virus]	89%	5.00E-70	47%	YP_008052687.1	
LMQMSMvls1241	Rep [Bat circovirus]	90%	1.00E-57	61%	AIF76274.1	
LMQMLGvls3293	hypothetical protein [uncultured marine virus]	51%	2.00E-32	35%	AGA18423.1	
LMQMLGvls3301	hypothetical protein [uncultured marine virus]	78%	4.00E-90	50%	AGA18377.1	
LMQMLGvls8177	Rep domain protein [Phaeocystis globosa virus]	98%	1.00E-43	47%	YP_008052687.1	
Virus Like Sequence	**2nd Closest BLASTx HIT**	**Query Cover**	**e-value**	**% identity**	**Accession #**	
LMSS2vls625	hypothetical protein [uncultured marine virus]	91%	3.00E-61	44%	AGA18407.1	
LMQMSMvls1241	replication-associated protein [Dragonfly larvae associated circular virus-2]	88%	2.00E-57	62%	YP_009001739.1	
LMQMLGvls3293	replicase [Cyclovirus sp.]	52%	3.00E-28	34%	ARO38325.1	
LMQMLGvls3301	hypothetical protein [uncultured marine virus]	77%	2.00E-83	49%	AGA18391.1	
LMQMLGvls8177	replication protein, partial [uncultured marine virus]	96%	3.00E-39	45%	GAC77766.1	
Virus Like Sequence	**1st Relevant Freshwater Virus BLASTx HIT**	**Query Cover**	**e-value**	**% identity**	**Accession #**	**Place in List**
LMSS2vls625	replication associated protein [Lake Sarah-associated circular virus-19]	85%	8.00E-43	39%	ALE29650.1	6
LMQMSMvls1241	replication associated protein [Lake Sarah-associated circular virus-41]	88%	1.00E-44	54%	YP_009237565.1	11
LMQMLGvls3293	replication-associated protein [Diporeia sp. associated circular virus]	50%	4.00E-26	35%	AGG39813.1	10
LMQMLGvls3301	replication-associated protein [Diporeia sp. associated circular virus]	67%	1.00E-57	46%	AGG39826.1	6
LMQMLGvls8177	replication associated protein [Lake Sarah-associated circular virus-19]	98%	4.00E-24	36%	YP_009237552.1	11

Table 3: CRESS-DNA virus-like sequences identified in my metavirome libraries with reads found in both the quagga mussel and sediment core libraries. Included in the table: the source library where the virus was identified, the contig number, the reference length of the sequence (nt), the number of reads found in each library, and finally information including query cover, e-value, % identity, and Accession # of the top BLASTx hit of each of CRESS DNA virus-like sequences (NCBI).

Source Library	Contig	Reference Length (nt)	LMQMS reads	LMQML reads	SS2 reads	SS4 reads	Total reads	Top BLASTx HIT	Query Cover	e-value	% identity	Accession #
SS2	732	348	0	1	3	0	4	hypothetical protein [uncultured marine virus]	99%	1.00E-55	72%	AGA18478.1
SS2	509	871	0	9	44	0	53	replication-associated protein [Sewage-associated circular DNA molecule]	81%	4.00E-89	56%	AJD07568.1
LMQML	8177	488	0	7	4	183464	183475	Rep domain protein [Phaeocystis globosa virus]	98%	1.00E-43	47%	YP_008052687.1
LMQML	8145	664	0	69	1	0	70	replication-associated protein [Sewage-associated circular DNA molecule]	61%	1.00E-25	53%	AJD07568.1
LMQML	7078	619	1	9	79	1	90	replication-associated protein [Sewage-associated circular DNA molecule]	60%	1.00E-37	52%	AJD07568.1
LMQML	4233	1390	0	39	5	0	44	replication associated protein [Lake Sarah-associated circular molecule 5]	75%	1.00E-151	75%	ALE29544.1
LMQML	3768	614	2	7	13	0	22	replication-associated protein [Sewage-associated circular DNA molecule]	62%	3.00E-19	52%	AJD07568.1
LMQML	3225	303	0	2	103	0	105	hypothetical protein [uncultured marine virus]	100%	1.00E-21	49%	AGA18388
LMQML	1077	536	0	6	52	0	58	replication-associated protein [Sewage-associated circular DNA molecule]	74%	3.00E-53	61%	AJD07568.1

Supplementary Material:

Methods:

LM29173 circovirus qPCR of sediment core layers

Quantitative PCR of the different sediment core layers in the three sediment cores was completed to quantify LM29173 circovirus load following the methods of Hewson et al. 2013a (Table S2; see extended methodology for more details). Two samples from each sediment core layer (two separate DNA extractions) were run in three experimental replicates. LM29173 circovirus qPCR analysis was completed for each of the eight sediment core subsamples in each of the 3 sediment cores (n= 48). Inhibition of the qPCR reaction was a problem in some of the sediment core layers and was corrected for (see extended methodology for more details).

Results:

LM29173 circoviruses in sediment cores

Quantitative PCR (qPCR) analysis results show that LM29173 circoviruses were detected throughout the nearshore and offshore sediment cores where *Diporeia spp.* were known to historically inhabit. Due to issues with qPCR reaction inhibition with all the sediment core subsamples, estimates of abundances of the LM29173 circovirus Rep genes could not be accurately determined. Even with high inhibition effects measured Ct values were above the no template controls indicating the presence of the LM29173 circoviruses throughout the sediment core subsamples (Figure S1).

Tables and Figures:

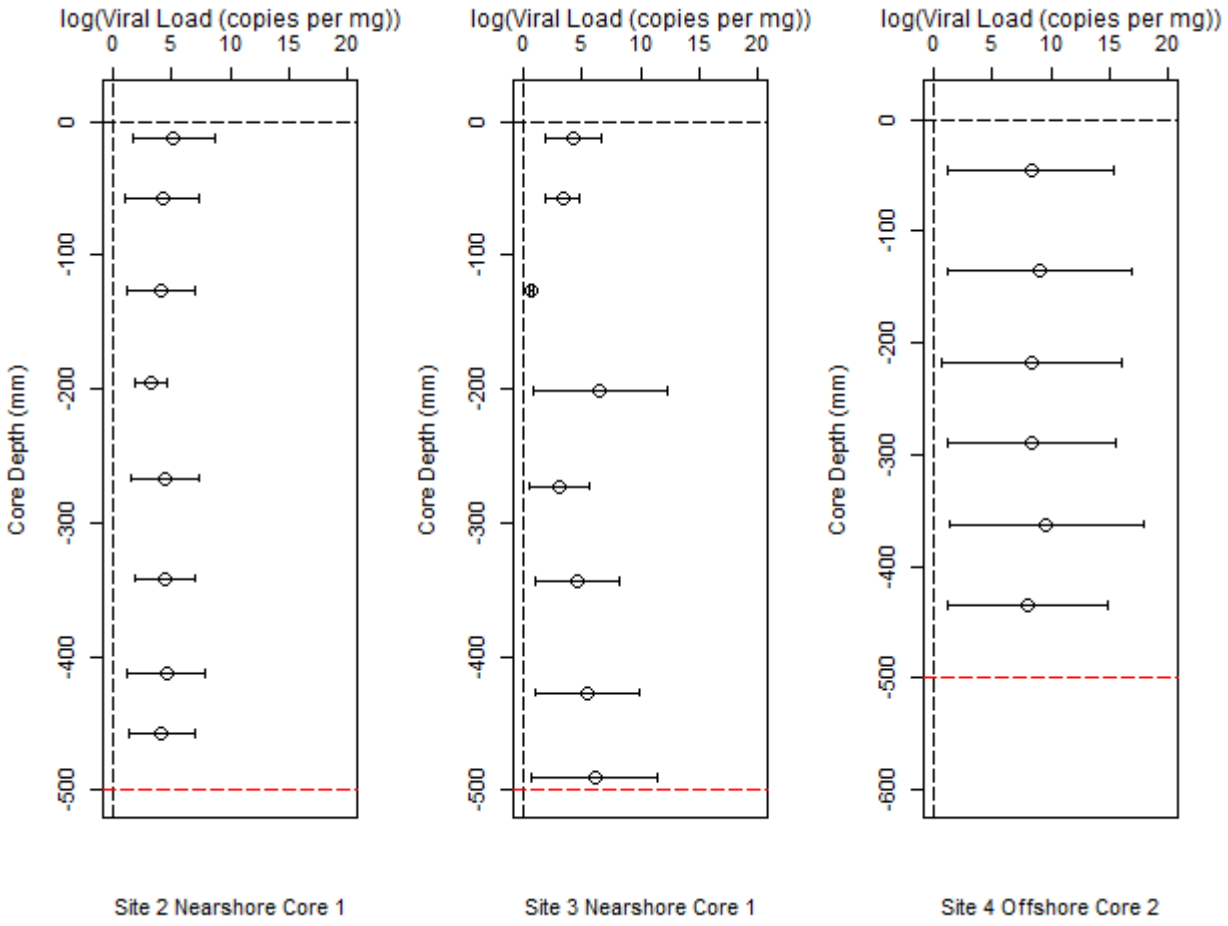


Figure S1: LM29173 circovirus qPCR analysis of the different sediment core layers. Far left= nearshore sediment core 1, middle = nearshore sediment core 2, and far right = offshore sediment core 2 (See Figure 1 for specific locations) The vertical axis represents depth in the sediment core in which the sample was taken with 0 being at the top of the core (=bottom of the lake) and-500 representing 500 millimeters down in the lake bottom. The horizontal axis represents the log transformation of the viral load (copies of rep gene of LM29173 per mg of sediment) in each of the sediment subsamples analyzed (SS1-SS8 with SS1 being at the top of the core). Viral load calculations were done with a correction for inhibition impacts in site 4 offshore core 2. The red dotted line indicated 50cm down in the sediment core.

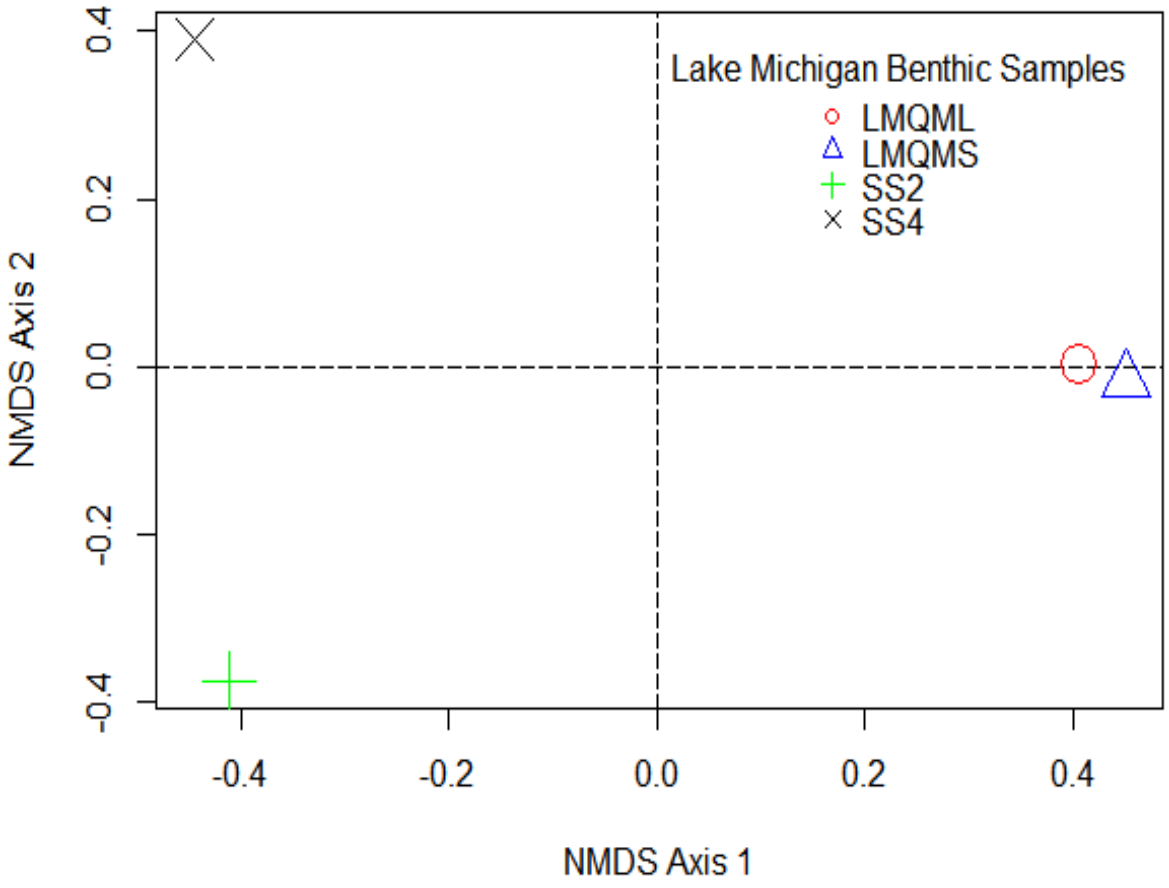


Figure S2: Nonmetric Dimensional Scaling (NMDS) ordination plot showing the differences in the viral like sequence reads found in the different metavirome library samples collected in the Lake Michigan benthos. LMQML= Lake Michigan Quagga Mussels Large (>24mm shell length) and is represented by the red circle. LMQMS = Lake Michigan Quagga Mussels Small (< 15mm shell length) and is represented by the blue triangle. SS2 stands = Subsample 2 (-136mm deep in the core) taken from offshore sediment core 4 and is represented by the green plus sign. SS4= Subsample 4 (-290mm deep in the core) taken from offshore sediment core 4 and is represented by the black X. The ordination plot was created using Bray-Curtis dissimilarity matrices in the isoNMDS function in the package VEGAN in Program R version 3.3.2 (R Core Development Team 2017).

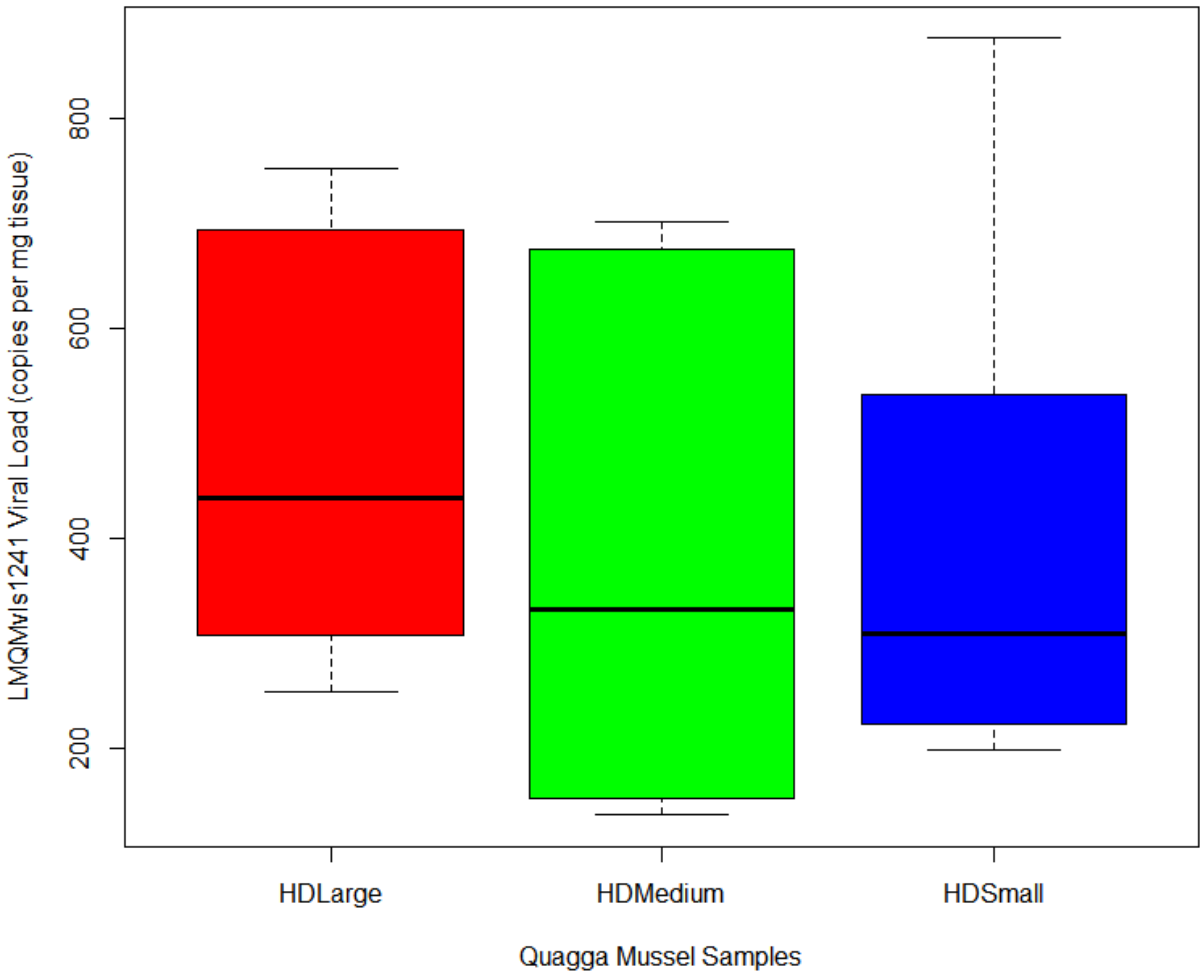


Figure S3: CRESS-DNA virus-like sequence 1241 qPCR analysis of different sized quagga mussels collected at the Holland Deep location (see Figure 1). The y-axis represents the average viral load (copies of LMQMLvls1241 Rep ORF copies per mg quagga mussel tissue) and the x-axis represents the samples large (n=2, >25 mm shell length), medium (n= 2, 15 mm-25 mm shell length), and small (n=2, <15 mm shell length) Lake Michigan quagga mussels.

Table S1: Quantitative PCR standards, probes, and primers for the two different CRESS-DNA virus/virus-like sequences analyzed in this study. Both probes used in this study had a FAM reporter dye and a TAMRA quencher dye and had the same thermocycling parameters. The standards, primers, and probes for LM29173 REP were the same as used in Hewson et al. 2013a. All qPCR reactions were run on a StepOnePlus Real Time PCR machine (Applied Biosystems).

Genotype	**Sequence**	**Thermocycling Parameters**
LMQMSvls1241	*Standard:* CTGATTATTCAAAAGCAGTTGGATTGGAGCGCAAGTGCATGGTCTTCTGGGGCAAAACTGGAACTGGTAAATCTAGACGTGCATGGGATGAGCCGGGTT *Probe:* [FAM] TGGTCTTCTGGGGCTAACTGCGA [TAMRA] *Forward-primer:* CCGGCTTCATCCCATGCAC *Reverse-primer:* GCAGTTGGATTGGAGCGCAAGTG	1 Cycle 95° C x 5min 60 cycles 95°C x 30s 58°C x 30s
LM29173 REP Hewson et al. 2013	*Standard:* AACGAGTTAGGGGAGTGTGGTACCCCACACATCCAGGCCTACGTTGCCTTGCTAAAGGACAAGAAGCAGAGGCAGGC *Probe:* [FAM] CCCCACACATCCAGGCC [TAMRA] *Forward-primer:* CGAGGTAGGGAGTGTGGTA *Reverse-primer*: TGCCTCTGCTTCTTGTCCTT	

Chapter 3

Extended Literature Review

Introduction:

The Great Lakes ecosystem has seen significant community changes as the result of nonindigenous species. The Lake Michigan benthic community in particular has been restructured from one dominated by the amphipod *Diporeia spp*. to one in which the invasive quagga mussels (*Dreissena rostriformis bugensis*) now dominates. This has had sweeping consequences for the entire Lake Michigan ecosystem including the microbial communities. The advent of metagenomic sequencing has revealed an important and not fully understood diversity of particular viruses, Circular Rep Encoding Single Stranded DNA viruses (CRESS-DNA) found in lake ecosystems. Recent studies reviewed CRESS-DNA viruses in *Diporeia spp*. found across the Great Lakes benthos. My thesis investigates the CRESS-DNA viral consortium found in the new invasive quagga mussel community in the Lake Michigan benthos and also those historically present in the sediments. The purpose of this review was to identify the relevant literature studies and questions that relate to my study of CRESS-DNA viruses found in quagga mussels and sediment cores in the central Lake Michigan benthos. The review is organized into two major sections: (1) a broad understanding of Great Lakes benthic ecology with a focus on Lake Michigan and (2) viruses in aquatic ecosystems with a focus on invertebrate CRESS-DNA viruses.

Great Lakes Benthic Ecology

Brief History

The Great Lakes ecosystems have changed profoundly in the last century as a result of the opening of the St. Lawrence Seaway to allow international trade to move into the Great Lakes. Construction of several canals that connected the Great Lakes watershed to the Atlantic Ocean allowed species from Atlantic coast to move in into the Great Lakes. With international trade came an influx of invasive species that moved in from the Atlantic coast and those from overseas that often hitched a ride via the ballast water of the transcontinental cargo ships. Jeff Alexander's book, "Pandora's Locks", provides an excellent overview of these changes (Alexander, 2009). Today there are more than 200 nonindigenous species in the Great Lakes basin (see https://www.glerl.noaa.gov/glansis/)

In Lake Michigan key nonindigenous species, including both those introduced by humans and those that were transported via transoceanic cargo vessels have created ecological changes throughout all parts of the ecosystem (e.g. top-down, middle-out, bottom up; Cuhel and Augilar, 2013). A review by Cuhul and Augilar (2013) describes the major changes in the Lake Michigan ecosystem caused by these key invasive species. The opening of the canals connecting the Atlantic Ocean to the Great Lakes allowed two key invaders to enter, the sea lamprey (*Petromyzan marinus*) and the alewife (*Alosa pseudoharengous*). Sea lampreys are a parasitic fish species that feeds on larger benthic game fish and when they arrived in the Great Lakes they caused the extirpation of the top predator fish, Lake Trout, resulting in top-down ecological change. The reductions in numbers of native predators resulted in open niche space for the next invader, the alewife, to thrive in the lake causing middle-out ecological change. The

alewife had massive population boom and bust cycles which resulted in mass die-offs of the fish to wash up on the Lake Michigan shores. This prompted action from the Michigan Department of Natural Resources (MDNR) and in 1966 they began to stock Pacific Salmon, starting with Coho in Lake Michigan, to feed on and control the alewife population (Tody and Tanner, 1966). This created an estimated billion dollar sports fishery for the salmon in Lake Michigan (Alexander, 2009). Following the establishment of the Pacific salmon in Lake Michigan, in the early 1990s invasive dreissenid mussels arrived on the scene where they established in the benthos and created bottom-up ecological changes through their filter feeding activity changing primary producer communities.

Dreissenid mussel invasion

The dreissenid mussels that have established in the Great Lakes include the zebra mussels (*Dreissenia polymorpha*) and quagga mussels (*Dreissenia rostriformis bugensis*). Both are native mussels to the Ponto-Caspian Sea region, with quagga mussels specifically from the Dnieper River drainage in Ukraine (Mills et al., 1993) and arrived in the Great Lakes via the ballast water of transoceanic cargo vessels (Nalepa and Schloessor, 1993; 2014). The zebra mussels came first, arriving in North America in the mid-1980s and the Great Lakes by the late 1980s (Hebert, 1989; Carlton, 2008). Zebra mussels quickly spread throughout the Great Lakes. Zebra mussels were able to thrive in the hard substrates of the littoral zones in all the Great Lakes except Lake Superior. In Lake Superior, calcium concentration and temperature were too low to support large populations of zebra mussels (Mellina and Rasmussen, 1994). Zebra mussels were able to fill an empty niche in the other Great Lakes attaching to a variety of rock, sand substrates as well as aquatic macrophytes in the littoral zone (Vanderploeg et al., 2002).

Quagga mussels (*Dreissenia rostriformis bugensis*) arrived and established after zebra mussels. The speculated delay in spread being due to zebra mussels sticking tighter to boat hulls. Also the habitats where the mussels first arrived at were shallow areas of lakes and rivers which the zebra mussels prefer compared to quiet deeper waters the quagga mussels prefer (Karatayev et al., 2015). Both zebra and quagga mussels have rapid filter feeding and reproduction physiologies that helped in their ability to effectively invade the Great Lakes (Vanderploeg et al., 2002). Dreissenid mussels are extremely effective at filtering water per unit mass. They also can filter a broad size range of particles (1 µm to 750 µm) and assimilate up to 40% of their body mass carbon each day (Vanderploeg et al., 2001). Dreissenid mussels have external fertilization with females producing as many as one million eggs and males 100 billion sperm producing planktonic offspring (Nalepa and Schloesser, 2014). These offspring move naturally in lake currents until they find a suitable habitat where they will attach themselves to become sessile adults.

Focusing on Lake Michigan, quagga mussels were the most successful and have become the dominant invertebrate in the Lake Michigan benthos. The review by Cuhel and Aguliar (2013) describes three reasons why quagga mussels were so successful in Lake Michigan. The first is the mussels' ability to actively feed at cold temperature ranging from 0.5°C to 5°C found in the deep waters of Lake Michigan and in the winter (Baldwin et al., 2002). Second is their lower metabolic rate which allows them to feed at these colder temperatures and still survive (Stockemann, 2003). Quagga mussels also have a lower temperature limit for reproduction than zebra mussels (Roe and MacIsaac, 1997). Third is quagga mussels are able to colonize soft substrates found in the offshore benthos. Quagga mussels even have a specific morphotype,

known as the profunda morph, characterized by light shell pigmentation and elongated shell and siphon that is adapted to burrowing in the soft substrate of cold deep waters (Dermott and Munawar, 1993; Claxton et al., 1998).

The establishment of the dreissenid mussels has had a significant impact on the entire Lake Michigan ecosystem. The largest impact comes from the filter feeding of the mussels which because of the large population densities that the mussels have established in the Great Lakes, take in large volumes of water quickly. This results in increases of water clarity from decreasing phytoplankton, seston, and organic matter which allows light to penetrate deeper into the water column (Karatayev et al., 2015). The filter feeding results in decreases in both phytoplankton and zooplankton density and diversity (Karateyev et al., 2015). Toxic algae blooms of the cyanobacteria *microcystsis* that have become problematic across the Great Lakes have been associated with the selective filter feeding of dreissenid mussels as well as the release of nutrients necessary for the blooms (Bykova et al., 2006). The impact of the dreissenid mussels differs across the different lakes and spatially across the littoral, profundal, and benthic zones. Quagga mussels in the deeper parts of the lake have the largest impact on spring diatom blooms, causing an overall decrease, during isothermal periods (Nalepa et al., 2010). In the offshore benthos quagga mussels have caused declines to most of the native organisms of these sediments, including amphipods (*Dipoeria spp*. which will be discussed in great detail in the next section), native molluscs, oligochaetes, and chironomids (Karateyev et al., 2015). Reductions are likely caused via competition (both direct and indirect) with quagga mussels for resources (e.g. algal and diatom particulates) and space with the quagga mussels. This has had impacts on important Lake Michigan fish species including the commercially important

coronoids (*Coregonus*; Whitefish and Bloaters) which have seen negative impacts on condition and population declines related to changing benthic food resources. Overall the mussels are causing benthification, the diversion of carbon from the lake to the benthos and also oligotrophication in which nutrients becoming limited reducing the productivity of the Lake Michigan ecosystem as a whole (Cuhel and Augilar, 2013).

Very few studies have looked at the impacts of dreissenid mussels have on microbial communities. A study by Lee et al. (2015) ran a series of microcosms with quagga mussels and sediments from Lake Michigan and found the presence of the quagga mussels increased sediment bacterial diversity and nitrifying bacterial taxa abundance. Thus the mussels are likely impacting the microbial aspects of nutrient cycling in Lake Michigan. Another part of the microbial communities includes viruses and one study found that zebra mussels could accumulate a low pathogenic avian influenza virus in their tissue (Strumpf et al., 2010). To my knowledge there have been no studies looking at viruses associated with quagga mussels in the Great Lakes.

Diporeia spp. decline

Diporeia spp. are a glacial relict benthic amphipod found in the Great Lakes region (Bousfield, 1989). They were originally named *Ponteporeia affinis* and were historically one of the most prevalent members of the benthic community in the upper Great Lakes (specifically Lake Michigan) in terms of biomass with estimates of *Diporeia* comprising greater than 70% of the biomass in offshore waters (>30m in depth; Eggleton, 1936; Cook and Johnson, 1974; Nalepa, 1989). *Diporeia spp*. are a burrowing amphipod that live in the colder deeper waters of the Great Lakes and feed on organic material that falls into the benthos from the pelagic zone.

A study by Marzolf (1965) showed that *Diporeia spp.* (still called *Ponteporeia* in this study) had a preference for finer organic sediment types with abundant bacteria present for the *Diporeia* to consume. Another study by Sly and Christie (1992) showed that *Diporeia* density was positively associated with organic carbon content in the sediment as well as sediment associated bacteria. *Diporeia* were an important food resource for many fish species including the commercially important whitefish and bloater (*Coregonus hoyi* and *Coregonsus clupeaformis, respectively*), slimy and deepwater sculpin (Cottus cognatus and Mysococephalus thompsonii, respectively), the nonindigenous baitfish alewifes and rainbow smelt (Alosa pseduoharengus and Osmerus mordax respectively), and yellow perch (*Perca flavseces*) (Anderson and Smith, 1971; Wells, 1980). Thus *Diporeia* help to recycle energy from primary productivity back to higher trophic levels (Nalepa et al., 2006)

Major declines of *Diporeia* populations in the Great Lakes started in the early 1990s documented first in eastern Lake Erie (Demott and Kerec, 1997), then eastern Lake Ontario (2001), southern Lake Michigan (Nalepa et al., 1998) and Saginaw Bay on Lake Huron (Nalepa et al., 2003). The declines followed the establishment of invasive dreissenid mussels in the lakes. There is consensus that the mussels played a definite role in *Diporeia* decline but the exact mechanism is still unknown (Nalepa et al., 2006). The major theories are that dreissenid mussels directly or indirectly influenced *Diporeia* food inputs (competition for spring diatom blooms), or that a concentrated toxin from dreissenid psuedofeces (undigested filter-feeding waste) impacted them, or they suffered from disease, but all of these theories have holes in them (Nalepa et al., 2006). In two cases in Lake Michigan and Lake Ontario, *Diporeia* populations started to decline before the establishment of dreissenids (Nalepa et al., 2009;

Watkins et al., 2007). In the Finger Lakes region of New York studies have found that *Diporeia* and dreissenids successfully coexist (Watkins et al., 2012). Both field and laboratory studies looking at the role of dreissenid mussels influencing *Diporeia* food resources have been inconclusive (Nalepa et al., 2006; Watkins et al., 2012). Studies looking at eukaryotic parasites and potential pathogenic bacteria associated with *Diporeia* have also been inconclusive with no real pattern in parasite load or bacterial distribution (Winters et al., 2014; 2015; Cave and Strychar, 2015). A study looking at dreissenid mussel pseudofeces, particularly sediments and *Diporeia* in mesocosms found they did not contain toxins although sediments exposed to dying *Diporeia* increased the mortality of other *Diporeia* placed in this same sediment indicating the presence of a potential pathogen (Laundrom et al., 2000; Dermott et al., 2005). This led to the work of Hewson et al. (2013a) who explored the possibility of a viral pathogen that may have played a role in *Diporeia* decline. They used a metagenomic approach (Thurber et al., 2009) to characterize the viruses present in different populations of *Diporeia* across the Great Lakes, finding a particular circovirus (a type of CRESS-DNA virus which will be discussed in greater detail later in the review) to be highly prevalent in the declining Lake Michigan *Diporeia* populations. Follow up studies have recently shown this virus to have little role in *Diporeia* physiology and that it is likely not a pathogen (Bistolas et al., 2017a; 2017b). Today *Diporeia* have been mostly extirpated from all the Great Lakes except Superior although a few deep water refuge populations in Lake Michigan and Lake Huron persist (Dr. Ashley Baldridge personal communication) and the mechanism for their declines remains unresolved.

Viruses in Aquatic Ecosystems:

Marine:

A review by Suttle (2007) describes all the major research that has been conducted involving marine viruses. The review indicated that the field of marine virology first made strides when high-throughput methods using epifluorescent microscopy could effectively quantify all the virus-like particles found in seawater (Noble and Fuhrman, 1998; Fuhrman, 1999). Virus like particles were found to be the most abundant component of seawater and are one of the largest pools of unknown genetic diversity on the planet. Most of the viruses found in marine systems are viruses of bacteria (bacteriophages). The first studies looked at the role of viruses in the microbial ecosystems in the oceans and have found that viruses play a key role in global biogeochemical cycling. Virus induced mortality in bacteria diverts a significant amount of carbon and nutrients away from the food web and organic matter pool in what is known as the viral shunt . A study by Wilhlem and Suttle (1999) estimates that a quarter of the primary production in the ocean ultimately flows through the viral shunt. This movement of carbon through the viral shunt influences the pool of available carbon and nutrients for food webs and also effects carbon sequestration in the ocean (Jover et al., 2014). Viruses influence bacterial mortality but estimating rates of viral lysis and the effect on host population is quite difficult. A key obstacle is adequate methodology to accurately estimate rates of viral mortality to include in models of global energy and nutrient cycling in the oceans (Suttle, 2007).

Viruses in marine ecosystems also help to structure microbial communities. It is believed viruses help control microbial communities via a mechanism called 'killing the winner' where a bacterial species that becomes abundant is put in check by virus infections allowing

other microbial species to take a foothold in the community (Murray and Jackson, 1992; Thingstad, 2000). This allows for a biodiverse community of microbes to flourish. These effects of viruses on microbial communities are variable along temporal and spatial gradients (Suttle, 2007). The community of viruses found in microbial communities and many other organisms is also quite diverse. New methods have been developed in order to investigate viral diversity and communities including shotgun metagenomic sequencing (Thurber et al., 2009). This genomic tool allows for the analysis of uncultured viral sequences to be analyzed from tissues and environmental samples (Thurber et al., 2009). The technique compares all DNA sequences in a sample to all virus-like sequences known in the sequence database GENBANK (NCBI) allowing for characterization of part of the viral community (most sequences are unknown and have no matches in the database) in a sample including novel sequences.

Freshwater:

In 2008, the Journal of Freshwater Biology ran a special issue to highlight the growth in studies of viruses in freshwater ecosystems. Wilhelm and Matteson (2008) published a review of the similarities and differences between freshwater and marine virioplankton. Studies in both marine and freshwater systems have yielded similar results with viruses in both systems regulating carbon and nutrient cycling and shaping microbial communities. Differences were seen with seasonal dynamics playing a larger role in freshwater systems, especially in lakes and rivers with distinct seasonal cycles (Lymer et al., 2008). Viral abundance was found to be more variable in a freshwater stream and a lake in Tennessee across the different seasons then compared to marine systems. Viral densities were found to be highest in the spring and lowest in the winter which corresponds with bacterial abundances (Lymer et al., 2008). Like marine

systems the most abundant virus in freshwater systems are those infecting bacteria (bacteriophages). Another difference provided in the review is that genetic studies have shown that viruses of marine and freshwater ecosystems are evolutionarily distinct.

Viruses can be found in almost all freshwater systems on this planet. As a huge source of unknown genetic diversity they offer a new frontier of discover. For example, a study published in the journal *Science* in 2009 showed a very high diversity in the viral community found in an Antarctic Lake that is frozen for most of the year (Lopez- Bueno et al., 2009). The study showed the lake to have the highest number of viral families found to date in a freshwater aquatic ecosystem. A unique issue that was brought up in the paper that is common in many viral diversity studies is the amount of unknown sequences that have no similarities with any sequences found in GenBank (NCBI). This study reported 86.7% of the sequences they found in the lake were unknown. This highlights just how little we actually know about what viruses and microbes are out there in the world, especially in aquatic systems.

CRESS-DNA viruses

As shown there is staggering amount of diversity in the viruses present in aquatic ecosystems, however the focus of my study was on an important group of viruses infecting eukaryotes known as Circular Rep Encoding Single Stranded (CRESS) DNA viruses. CRESS-DNA viruses consist of a simple circular single stranded genome architecture usually smaller than 6kb coding for several proteins including a replication initiator protein (Rep) that initiates the replication of all other genetic components of the virus (Rosario et al., 2008). CRESS-DNA viruses have been discovered in wide range of aquatic environments, from the open ocean (Breitbart et al., 2002; Rohwer and Thurber, 2009; Ng et al., 2013), deep sea vents and

sediments (Yoshinda et al., 2013), Antarctic lakes (Lopez-Bueno et al., 2009), sewage treatment ponds (Kraberger et al., 2015) and freshwater lakes (Dayaram et al., 2013; 2015; Hewson et al., 2013; Bistolas et al., 2017). Recent studies have found CRESS-DNA viruses to be associated with aquatic invertebrates including amphipods (Hewson et al., 2013; Bistolas et al., 2017; In Press), copepods (Dunlap et al., 2013), molluscs (Dayaram et al., 2013; 2015; 2016), insect larva (Dayaram et al., 2016), and echinoderms (Jackson et al., 2016). The use of metaviromic sequencing (Thurber et al., 2009) with the particular step of rolling circle amplification (RCA) allows for access to the rare species sequences in the environment has a known bias towards CRESS-DNA viral elements with their primarily circular genomes (Binga et al., 2008; Rosario et al., 2012). This permitted the discovery of these viruses across aquatic ecosystems.

With the discovery of the large diversity of CRESS-DNA viruses in aquatic environments ecological and evolutionary questions about what exactly these viruses are doing in these aquatic ecosystems are being investigated. In invertebrates in particular a wide diversity of these viruses have been detected in association with invertebrate species based on metaviromic and qPCR viral load results; however only a few have documented microscopic evidence of the CRESS-DNA viral infection (Dunlap et al., 2012; Hewson et al., 2013b) and even in these studies it cannot be directly proven that the virus like particles they observed in the tissues of the invertebrates are in fact the CRESS-DNA viruses they investigated. An extensive study by Hewson et al. (2013b) identified a particular CRESS-DNA virus and a single stranded RNA virus via metagenomics that is associated with important *Daphnia spp.* in a freshwater lake. The study used qPCR to follow the dynamics of viral infection in the *Daphnia* spp. over a summer season showing that viral prevalence increases before major population declines of

Daphnia in the lake (Hewson et al., 2013b). Recent work looking at the amphipod *Diporeia* in the Great Lakes region were unable to show any physiological cost of associated CRESS-DNA viruses (Bistolas et al., 2017a) and no clear patterns in gene expression (e.g. immune response etc.) in *Diporeia* infected with high loads of CRESS-DNA viruses (Bistolas et al., 2017b) indicating a limited role in these viruses in major population declines of *Diporeia* spp. in the Great Lakes. This supports the hypothesis that many of these CRESS-DNA viruses may be persistent and asymptomatic infections (Okamoto, 2009; Brajea de Oliveria, 2015; Bistolas et al., 2017b).

Many questions still remain about how CRESS-DNA viruses move between organisms in aquatic ecosystems, where they originate, and how they behave in the environment (e.g. sediments). An extensive study characterized CRESS-DNA viruses sampled from a diversity of different organisms and habitats in a freshwater lake in New Zealand (Dayaram et al., 2016) in an attempt to understand how CRESS-DNA viruses are related based on feeding relationships. The CRESS-DNA viruses they detected were widespread across the organism and habitats studied with molluscs and bivalves showing a high number of CRESS-DNA viruses likely via bioaccumulation via filter feeding and grazing activities (Dayaram et al., 2016). Bivalves are well known for their ability to bioaccumulate human and other vertebrate viruses in their tissues via this filter feeding (Elston, 1997) and Dayaram et al. (2016) suggests that they may be a good target for biomonitoring of CRESS-DNA and other viruses in aquatic ecosystems. Due to this bioaccumulation of viruses it is difficult to determine which viruses are actually infecting the bivalves versus those that are infecting other host species that bivalves may be consuming. With many of the metavirome studies looking at CRESS-DNA viruses in invertebrates they have

been shown to be associated with these invertebrates but direct evidence for their hosts has not been shown (e.g. proving Koch's postulates using microscopy and pathology). This leads to the important hypothesis that much of CRESS-DNA viruses observed associated with invertebrates may be infecting eukaryotic algae and fungi which are highly diverse and abundant in freshwater aquatic ecosystems (Rosario et al., 2012; Dayaram et al., 2016). Further investigations into what exactly these viruses are infecting will help to inform answers about their ecology, origins, and evolution.

CRESS-DNA viruses in sediments

DNA from sediments cores collected in temperate lakes has been used to look back at historic ecosystems of cyanobacteria (Pal et al., 2015). Sediments in aquatic ecosystems are natural reservoirs for many viruses including those from the pelagic zone that sorb to suspended solids and sink to the benthic sediment and those of benthic origin (Hewson et al., 2001; Hewson and Fuhrman, 2003). These viruses may remain intact and persist deep within the anoxic sediments for long time periods (Suttle, 2000; Lawerence et al., 2002; Coolen, 2011; Brown unpublished data). This allows for a repository of past viral populations present in a lake permits important investigations in origins and evolution of different viruses (Brown unpublished data). An important study looking at sediment cores from a freshwater lake using metaviromics to characterize cyanophage viral populations present found both persistent genotypes that have always been present in the lake and those that have recently dispersed into the lakes (Brown unpublished data). To my knowledge there has been no research looking into the CRESS-DNA viruses preserved in benthic sediment cores associated with aquatic benthic and pelagic invertebrates in a lake ecosystem.

Conclusions and Future Directions:

Invasive species have had a tremendous impact on the Lake Michigan ecosystem. Today the quagga mussels are the dominant player in the benthic community and their impacts are felt system wide. With the use of metagenomic sequencing, investigations into the once hidden realm of viral diversity and ecology can be investigated. A particular group of viruses, the CRESS-DNA viruses were found to be important in the mostly extirpated amphipod *Diporeia* in the Lake Michigan benthos. Nothing else is known about CRESS-DNA viruses in the Lake Michigan benthos. In the larger context of freshwater ecology many questions remain about these viruses in lake ecosystems about what their role is, what they infect, how they're moved, what their origins are, and how they're preserved etc. Future research should look to close these gaps, to clarify what, if any, significant role these viruses may play in large lake ecosystems and how invasive species impacts may be seen from the large ecological processes all the way into the small world of viral ecology.

Extended Methodology:

Rearing of quagga mussel in the lab

Live quagga mussels collected at the Holland Deep location (Figure 1) were housed in aquaria with other mussels of the same size class. The aquaria were approximately 75 liters in volume and were filled with filtered tap water from the lab sink to an approximately 10 cm height. A bubbler was used to provide aeration. Conditions were quite artificial with no natural lake sediment, natural light regime, or temperature regime (which would be dark and cold conditions of the deep offshore waters). The mussels were fed Kent Zoecon phytoplankton food weekly from the end of June to the end of December 2016 prior to viral analysis. The mussels were removed from the aquaria and immediately frozen in liquid nitrogen and then placed in the -80 °C freezer. Samples of pseudofeces from the bottom of tank were also taken, flash frozen in liquid nitrogen, and placed in the -80°C freezer.

LM29173 circovirus qPCR of sediment core layers

Quantitative PCR of the LM29173 circovirus rep-gene was done following the methods of Hewson et al. (2013a). The reactions were run on an Applied Biosystems Real Time PCR instrument. Three replications of 25 µl qPCR reactions (5 µl of extracted DNA in each reaction) were run with 1X TaqMan Universal Master Mix II with no UNG (ThermoFisher Scientific, Waltham, MA, USA), 200 pmol of each of the forward (Table 5; LM29173_F; 5'-CHAGGTAGGGGAGTGTGGTA-3') and reverse (Table S1; LM29173_R; 5'-TGCCTCTGCTTCTTGTCCTT) primers and probe (Table S1; LM29173_Pr; 5'-CCCCACACATCCAGGCC-3') labeled at the 5' with 6-FAM. The thermal cycling was 95°C for 5 min for 60 cycles of denature (95°C for 3 sec) and anneal (58°C for 30 sec). An oligonucleotide

standard of the LM29173 rep gene (Table S1; LM29173Std; 5'-AACGAGTTAGGGGAGTGtGGTACCCCACACATCCAGGCCTACGTTGCCTTGCTAAAGGACAAGAAGCAGAGGCAGGC-3') was created using the program PRIMER3 (Rozen and Skaletsky, 2000) and was included in each qPCR run in an eight fold dilution from 10^8 to 10^1 copies. The standard curve and threshold cycle (C_t) was determined using the AB1 7300 software (Applied Biosystems) using a least squares R^2 value. The output files after a qPCR run were transferred to a MS excel (2016) file for further analysis. The genotype copy number was multiplied by 2 to account for the viruses that are single stranded and the qPCR products that are double stranded. In each qPCR run for each sample reaction inhibition was examined by running a sample with an added aliquot of 10^4 standard. Inhibition may be caused by any agents that disrupt the qPCR reaction not removed during the DNA extraction procedure. Samples were considered inhibited when the spiked samples measured to have less than 10^4 copies. Since the amount of sample DNA was the same in each qPCR reaction (5 µl), samples that showed inhibition were corrected by taking the difference of 10^4 and the measured value in the spiked sample and adding this value to each to the measured value of each of the other replications. A sample was considered to have the viruses present if all the replications had a value of greater than 10 copies and there was not a >1 magnitude difference between the replications (Hewson et al., 2013). Viral load was measured as the number of viral genotype copies per mg of tissue or sediment.

Metavirome Data Analysis

The CRESS-DNA virus-like sequence reads found in each library were analyzed using non-metric dimensional scaling (NMDS; Clarke, 1993) in order to visualize differences in the CRESS-DNA virus-like sequence consortia found in each sample library. NMDS ordination was

completed using a Bray-Curtis dissimilarity matrix in the isoNMDS function in the package VEGAN in Program R version 3.3.2 (R Core Development Team, 2017). Other dissimilarity matrices were tried including Euclidean and Manhattan with the same output visualization for each.

LMQMvls1241 qPCR standard and primer design

LMQMvls1241 was chosen for further investigation using qPCR analysis. BLASTx analysis showed good alignment throughout the sequence with known CRESS-DNA virus rep genes (NCBI). The program Primer3 v 0.4.0 was used to design and find the best locations for creation of a standard, primers, and probes in LMQMvls1241 (v 0.4.0 Rozen and Staletsky 2000). The output showed the best location for primer attachment to be at the end of the sequence from 490bp to 574bp to create an 84bp product. Thus I designed a 100bp standard sequence from 476bp to 576bp in the LMQMvls1241 sequence which was created by IDT (Integrated DNA Technology; Skokie ILL, USA). Further analysis using ORFinder (NCBI) showed several open reading frames throughout the virus-like sequence including ORF3 which was found at the end of the sequence were the standard was created. ORF1 was found in the beginning of the sequence and was found to have to align with the Viral Rep Superfamily motif. This analysis was done after considerable effort in creating and testing the standard and primers and in hind sight, it would have been best to create a standard within ORF1 that contained the best alignment with a viral Rep gene. LMQMvls1241 represents a small incomplete viral genome and the standard we created still contained a partial ORF still representative of the virus-like sequence we were trying to quantify.

The LMQMvls1241 standard created was used to test primers via PCR reaction. Upon arriving to the lab the standard was reconstituted with Milli-Q water in a sterile fume hood to avoid contamination. I ordered a set of four primers based on the top four results in Primer3 (v 0.4.0 Rozen and Staletsky 2000) and tested all of these on the standard to see which produced the correct product with good yield. For PCR reaction setups 10 µl reactions were completed with the following reaction cocktail; 4 µl Milli-Q water, 1 µl 10X PCR buffer [200 mM Tris HCl (pH 8.4), 500 mM KCl], 0.8µl 25 mM Mg^{2+} ,0.2µl 10mM dNTP Mix (dATP, dCTP, dGTP and dTTP, each at a final concentration of 10 mM), 0.8 µl 5 µM Forward Primer, 0.8 µl 5 µM Reverse Primer, 0.4 µl 5U/µl Taq Polymerase, and 2µl LMQMvls1241 template. All reagents were from ThermoFisher Scientifc, Waltham, MA, USA. PCR reactions run in triplicate were completed in an Eppendorf Mastercycler gradient with a 95 °C denaturing step for 5min minutes followed by 30 cycles of 58° C annealing step for 30 seconds and then a 95°C denaturing step for 30 seconds. PCR products were visualized using Gel Electrophoresis run in a 2% sucrose gel agar at 90V for 45 minutes to 90 minutes and then placed in a SYBR Gold stain solution added after the completion of gel for 30 minutes. Pictures of the gel were taken in a UV visualizer and successful products appeared with strong bands on the plate between the 75bp and 100bp ladder (84bp product). Primer set 1 was found to work the best and was used for qPCR (Table 5). A hybridization probe was designed in Primer3 (v 0.4.0) also to bind between the two primer attachments points to add specificity to the qPCR reaction (v 0.4.0 Rozen and Staletsky 2000). The specialized hybridization probe was created by IDT with a TAMRA quencher dye and FAM reporter dye (Integrated DNA Technology; Skokie ILL, USA).

References:

Alexander, J., 2009. Pandora's locks: The opening of the Great Lakes-St Lawrence Seaway. East Lansing, MI: Michigan State University Press, 431p.

Anderson, E.D., Smith, L.L. Jr., 1971. A synoptic study of food habits of 30 fish species from western Lake Superior. University of Minnesota Agricultural Expo Student Technical Bulletin 279.

Baldwin, B.S., Mayer, M.S., Dayton, J., Pau, N., Mendilla, J., Sullivan, M., Moore, A., Ma, A., Mills, E.L., 2002. Comparative growth and feeding in zebra and quagga mussels (Dreissena polymorpha and Dreissena bugensis): implications for North American Lakes. Canadian Journal of Fisheries and Aquatic Science 59, 680-694.

Barbiero, R.P., Schmude, K., Lesht, B.M., Riseng, C.M., Warren, G.J., Tuchman, M.L., 2011. Trends in *Diporeia* populations across the Laurentian Great Lakes, 1997-2009. Journal of Great Lakes Research 37, 9-17.

Binga, E.K., Lasken, R. S., Neufeld, J.D., 2008. Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. International Society of Microbial Ecology Journal 2, 233-241.

Bistolas, K.S.I., Jackson, E.W., Watkins, J.M., Rudstam, L.G., Hewson, I., 2017a. Distribution of circular single-stranded DNA viruses associated with benthic amphipods of genus *Diporeia* in the Laurentian Great Lakes. Freshwater Biology 62(7), 1220-1231.

Bistolas, K.S.I., Rudstam, L.G., Hewson, I., 2017b. Gene expression of benthic amphipods (genus: *Diporeia*) in relation to a circular ssDNA virus across two Laurentian Great Lakes. PeerJ e3810; DOI 10.7717/peerj.3810.

Bousfield, E.L., 1989. Revised morphological relationships within the amphipod genera Pontoporeia and Gammaracanthus and the "glacial relict" significance of their postglacial distributions. Canadian Journal of Fisheries and Aquatic Sciences 46, 1714-1725.

Brajao de Oliveria, K., 2015. Torque teno virus; a ubiquitous virus. Revista Brasileira de Hematolgia e Hemterpia 37(6), 357-358.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology 13, 278-284.

Brown, J., 2014. Comparison of water column virus populations over a 180+ year time period within the varved sediments of Fayetteville Green Lake. Chapter 4 of Dissertation. Cornell University.

Bykova, O., Laursen, V., Bostan, J., Bautista J., McCarthy L., 2006. Do zebra mussels (*Dreissena polymorpha*) alter lake water chemistry in a way that favors *Microcystis* growth? Science of the Total Environment 371, 362-372.

Carlton, J.T., 2008. The zebra mussel *Dreissena polymorpha* found in North America in 1986 and 1987. Journal of Great Lakes Research 34, 770-773.

Cave, C.S., Strychar, K. B. 2014. Decline of *Diporeia* in Lake Michigan: Was Disease Associated with Invasive Species the Primary Factor? International Journal of Biology 7, 93.

Cook, D.G., Johnson, G., 1974. Benthic macroinvertebrates of the St Lawrence Great Lakes. Journal of Fisheries Research Board of Canada 31, 763-782.

Coolen, M.J., 2011. 7000 years of Emiliania huxlyi viruses in the Black Sea. Science 33, 451-452.

Claxton, W.T., Wilson, A.B., Mackie, G.L., Boulding, E.G., 1998. A genetic and morphological comparison of shallow- and deep- water populations of the introduced dreissenid bivalve *Dreissena bugensis*. Canadian Journal of Zoology 76, 1269-1276.

Cuhel, R.L., Aguilar, C., 2013. Ecosystem Transformations of the Laurentian Great Lake Michigan by Nonindigenous Biological Invaders. Annuals Review in Marine Sciences 5, 289-320.

Dayaram, A., Goldstein, S., Zawar-Reza, P., Gomez, C., Harding, J.S., Varsani, A., 2013. Identification of starling circovirus in an estuarine mollusc (Amphibola crenata) in New Zealand using metagenomics approaches. Genome Announcement 1(3), e0027-13.

Dayaram, A, Goldstien, S., Arguello-Astroga, G.R., Zawar-Reza, P., Gomez, C., Harding, J.S., Varsani, A. 2015. Diverse small circular DNA viruses circulating amongst estuarine molluscs. Infections, Genetics and Evolution 31, 284-295.

Dayaram, A., Galtowitsch, M.L., Arguello-Astorga, G.R., Van Bystervelt, K., Kraberger, S. Stainton, D., Harding, J.S, Roumagnac, P., Martin, D.P. Lefeuvre, P., Varsani, A., 2016. Diverse circular replication-associated protein encoding viruses circulating within a lake ecosystem. Infection, Genetics and Evolution 39, 304-316.

Dermott, R., Munawar, M., Bonnell, R., Carrou, S., Nibblock, H., Nalepa, T., Messick, G., 2005. Preliminary investigations for the causes of the disappearance of *Diporeia spp*. from Lake Ontario. Pp. 203-232 in Great Lakes Fisheries Commission technical report. Great Lakes Fisheries Commission. Ann Arbor, MI.

Dermott, R., Munawar, M., 1993. Invasion of Lake Erie offshore sediments by Dreissena, and its ecological implications. Canadian Journal of Fisheries and Aquatic Sciences 50, 2298-2304.

Dermott, R., Kerec, D., 1997. Changes to the deep water benthos of eastern Lake Erie since the invasion of Dreissena: 1979-1993. Canadian Journal of Fisheries and Aquatic Sciences 50, 2298-2304.

Dunlap, D. S., Ng, T. F. F., Rosario, K., Barbosa, J. G., Greco, A. M., Breitbart, M., Hewson, I., 2013. Molecular and microscopic evidence of viruses in marine copepods. Proceedings of the National Academy of Sciences 110(4), 1375-1380.

Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. National Review of Microbiology 3, 504-510.

Elston, R., 1997. Special topic review: Bivalve mollusk viruses. World Journal of Microbiology and Biotechnology 13, 393-403.

Fitzgerald, S.A., Gardner, W.S., 1993. An algal carbon budget for pelagic-benthic coupling in Lake Michigan. Limnology and Oceanography 38(3), 547-560.

Fuhrman, J. A., 1999. Marine viruses and their biogeochemical and ecological effects. Nature, 399(6736), 541-548.

Gastrich, M.D., Leigh-Bell, J.A., Gobler, C.J., Anderson, O.R., Wilhelm, S.W., Bryan M., 2004. Viruses as potential regulators of regional brown tide blooms caused by the alga, *Aureococcusanophagefferens*. Estuaries 27(1), 112-119.

Hebert, P. D. N., Muncaster, B.W., Mackie, G.L., 1989. Ecological and Genetic Studies on Dreissenia polymorpha (Pallas): A New Mollusc in the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 46, 1587-1591.

Hewson, I., Barbosa, J.G., Brown J.M., Donelan, R.P., Eaglesham, J.B., Eggleston, E.M., LaBarre, B.A., 2012. Temporal dynamics and decay of putatively allochthonous and autochthonous viral genotypes in contrasting freshwater lakes. Applied and Environmental Microbiology 78(18), 6583-6591.

Hewson, I., Eaglesham, J.B., Hook, T.O., LaBarre, B.A., Sepulveda, M.A., Thompson, P.D., Watkins, J.M., Rudstam, L.G., 2013a. Investigation of viruses in *Diporeia spp.* from the Laurentian Great Lakes and Owasco Lake as potential stressors of declining populations. Journal of Great Lakes Research 39, 499-506.

Hewson, I., O'Neil J. A., Fuhrman J.A., Dennison, W.C., 2001. Virus-like particle distribution and abundance in sediments and overlying waters along eutrophication gradients in two subtropical estuaries. Limnology and Oceanography 46(7), 1734-1746.

Hewson, I., Fuhrman J.A., 2003. Viriobenthos production and virioplankton sorptive scavenging by suspended particles in coastal and pelagic waters. Microbiology Ecology 46, 337-347.

Hewson, I., Ng, G., Li, W., LaBarre, B.A., Aguirre, I., Barbosa, J.G., Breitbart, M., Greco, A.W., Kearns, C.M., Looi, Al, Schaffner, L.R., Thompson, P.D., Hairston, N.G., 2013b. Metagenomic identification, season dynamics and potential transmission mechanisms of a Daphnia- associated single- stranded DNA virus in two temperate lakes. Limnology and Oceanography 58(5), 1605-1620.

Hurst, C.J., 2000. Viral Ecology. Academic Press, 659p.

Jackson, E.W., Bistolas, K.S., Button, J.B., Hewson, I., 2016. Novel Circular Single-Stranded DNA Viruses among an Asteroid, Echinoid, and Holothurian (Phylum Echinodermata). PLoS ONE 11(11), e0166093.

Karatayev, A.Y., Burlakova, L.E., Padilla D.K., 2015. Zebra versus quagga mussels: a review of their spread, population dynamics, and ecosystems impacts. Hydrobiologia 746, 97-112.

Keller, R. P., Cadotte M. W., Sandiford, G., 2014. Invasive Species in a Globalized World. University of Chicago Press, 427p.

Landrum, P.F., Gossiaux, D.C., Nalepa, T. F., Fanslow, D. L., 2000. Evaluation of Lake Michigan Sediment for Causes of the Disappearance of *Diporeia spp.* in Southern Lake Michigan. Journal of Great Lakes Research 26(4), 402-407.

Lee, P.O., McLellan, S.L., Graham, L.E., Young E.B., 2015. Invasive dreissenid mussels and benthic algae in Lake Michigan: characterizing effects on sediment bacteria communities. FEMS Microbiology Ecology 91: 1-13.

Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., Alcami, A., 2009. High Diversity of the Viral Community from an Antarctic Lake. Science New Series 326, 858-861.

Lymer, D., Logue, J.B., Brussard, C.P.D., Baudoux, A.C., Vrede, K., Lindstrom, E.S., 2008. Temporal variation in freshwater viral and bacterial community composition. Freshwater Biology 53, 1163-1175.

Marzolf, R. G., 1965. Substrate Relations of the Burrowing Amphipod *Pontoporeia Affinis* in Lake Michigan. Ecology 46(5), 580-592.

Mellina, E., Rasmussen, J.B., 1994. Patterns in the distribution and abundance of zebra mussel (*Dreissena polymorpha*) in rivers and lakes in relation to substrate and physicochemical factors. Canadian Journal of Fisheries and Aquatic Sciences 51, 1024-1036.

Middleboe, M., Jacquet, S., Weinbauer, M., 2008. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. Freshwater Biology 53, 1069-1075.

Murray, A.G., Jackson, G.A., 1992. Viral dynamics: a model of the effects size, shape, motion and abundance of single-celled planktonic organisms and other particles. Marine Ecology Program Series 89, 103-116.

Nagasaki, K., Tarutani, K., Tomaru, Y., Itakura, S., Tamai, K., Yamanaka, S. Yamaguchi, M., 2002. Possible use of viruses as a microbiological agent against harmful algal blooms. Fisheries Science 68, 497-500.

Nalepa, T.F., Hartson, D.J., Fanslow, D.L., Lang, G.A., Lozano, S.J., 1998. Declines in benthic macroinvertebrate populations in southern Lake Michigan, 1980-1993. Canadian Journal of Fisheries and Aquatic Science 55, 2402-2413.

Nalepa, T.F., Fanslow, D.L, Iii, A.J.F., Lang, G.A., Eadie, B.J., Quigley, M.A., 2006. Continued disappearance of the benthic amphipod *Diporeia spp.* in Lake Michigan: is there evidence for food limitation? Canadian Journal of Fisheries and Aquatic Science 63, 872-890.

Nalepa, T.F., Fanslow, D.L., Lang, G.A., 2009. Transformation of the offshore benthic community in Lake Michigan: a recent shift from the native amphipod *Diporeia spp.* to the invasive mussel *Dreissena rostriformis bugensis.* Freshwater Biology 54, 466-479.

Nalepa, T.F., Schloesser, D.W., 1993. Zebra Mussels Biology, Impacts and Contol. Lewis Publishers, Boca Raton, 810p.

Nalepa, T.F., Schloesser, D.W., 2014. Quagga and Zebra Mussels: Biology, Impacts and Control. 2nd edition. CRC Press, Boca Raton, 7750p.

Noble, R.T., Fuhrman, J.A., 1998 .Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquatic Microbial Ecology 14, 113-118.

Ng, T.F.F., Willner, D., Nilsson, C., Lim, Y.W., Schmieder, R., Chau, B., Ruan, Y., Rohwer, F., Breitbart, M., 2010. Vector-based metagenomics for animal virus surveillance. International Journal Infectious Disease 14, E378.

Pagnucco, K.S., Maynard, G.A., Fera, S.A., Yan, N.D., Nalepa, T.F., Ricciardi A., 2015. The future of species invasions in the Great Lakes-St. Lawerence River basin. Journal of Great Lakes Research 41, 96-107.

Pal, S., Gregory-Eaves, I., Pick, F.R., 2015. Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. Journal of Paleolimnology 54, 87-101.

Rosario, K., Duffy, S., Breitbart, M., 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. Journal of General Virology 90, 2418-2424.

Rosario, K. Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic single-stranded DNA viruses: insights gained from metagenomics. Archives of Virology 157, 1851-1871.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and biologists programmers. In: Krawertz. S., Misener S. (Eds.), Bioinformatics Methods and Protocols, Humana Press., Totowa,NJ, pp. 365-386.

Sly, P.G., Christie, W.J., 1992. Factors influencing densities and distributions of *Pontoporeia hoyi* in Lake Ontario. Hydrobiologia 235, 321-352.

Stoeckmann, A., 2003. Physiological energetics of Lake Erie dreissenid mussels; a basis for the displacement of *Dreissena polymorpha* by *Dreissena bugensis*. Canadian Journal of Fisheries and Aquatic Sciences 60, 126-134.

Stumpf, P., Failing, K., Papp, T., Nazir, J., Bohm, R., Marschange, R.E., 2010. Accumulation of low pathogenic avian influenza virus in zebra mussels (*Dreissena polymorpha*). Avian Disease 54(4), 1183-1890.

Suttle, C.A., 2007. Marine viruses- major players in the global ecosystem. National Review of Microbiology 5, 801-812.

Suttle, C.A., 2013. Viruses: unlocking the greatest biodiversity on Earth. Genome 10, 542-544.

Tody, W.H., Tanner, H.A., 1966. Coho salmon for the Great Lakes. Fisheries Management Report 1. Michigan Department of Conservation Ann Arbor.

Thingstad, T.F., 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnology and Oceanography 45, 1320-1328.

Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. Nature protocols 4(4), 470-483.

Van Etten, J.L., Van Etten, C.H., Johnson, J.K., Burbank, D.E., 1985. A survey for viruses from freshwater that infect a eukaryotic Chlorella-like green alga. Applied Environment Microbiology 49, 1326-1328.

Vanderploeg, H.A., Liebig, J.R., Carmichaelm, W.M., Agy, M.A., Johengen, T.H., Fahnenstiel, G.F., Nalepa, T.F., 2001. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic Microcystsis blooms in Saginaw Bay (Lake Huron) and Lake Erie. Canadian Journal of Fisheries and Aquatic Sciences. 58, 1208-1221.

Vanderploeg, H.A, Nalepa, T.F., Jude, D.J., Mills, E.L., Holeck, K.T., Liebig, J.R., Grigorovich, I.A., Ojaveer H., 2002. Dispersal and emerging ecological impacts of Ponto- Caspian species in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 59, 1209-1228.

Vanderploeg, H.A. Liebid, J.R., Nalepa, T.F., Fahnenstiel, G.L., Pothoven, S.A., 2010. Dreissena and the disappearance of the spring phytoplankton bloom in Lake Michigan. Journal of Great Lakes Research 36: 50-59.

Watkins, J.M., Dermott, R., Lozano, S.J., Mills, E.L. Rudstam, L.G., Scharold, J.V., 2007. Evidence for remote effects of dreissenid mussels on the amphipod Diporeia: analysis of Lake Ontario Benthic Surveys, 1972-2003. Journal of Great Lakes Research 33, 642-657.

Wells, L., 1980. Food of alewives, yellow perch, spottail shiners, trout-perch, and slimy and fourhorn sculpins in southeastern Lake Michigan. U.S. Fish and Wildlife Service Technical Report, 98p.

Wilhlem, S.W., Suttle, C.A., 1999. Viruses and nutrient cycles in the sea. Bioscience 49, 781-788.

Winnell, M.H., White, D.S., 1984. Ecology of Shallow and Deep Water Populations of *Pontoporiea hoyi* (Smith) (Amphipoda) in Lake Michigan. Freshwater Invertebrate Ecology 3, 118-138.

Winters, A.D., Fitzgerald, S., Brendan T.O., Nalepa, T., Faisal, M., 2014. Spatio-temporal dynamics of parasites infecting *Diporeia spp.* (Amphipoda, Gammaridae) in southern Lake Michigan (USA). Journal of Invertebrate Pathology 121, 35-37.

Winters, A.D., Marsh, T.L., Brenden, T.O., Faisal, M. 2015. Analysis of bacterial communities associated with the benthic amphipod *Diporeia* in the Laurentian Great Lakes Basin. Canadian Journal of Microbiology 61, 72-81.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., Takai, K., 2013. Metagenomic Analysis of Viral Communities in (Hado) Pelagic Sediments. PLoS One 8(2), e5727.