8-2024

Investigation of the Microbiome of a Recirculating Aquaculture System (RAS) for Crassostrea virginica Larvae

Isabelle Townsend

INVESTIGATION OF THE MICROBIOME OF A RECIRCULATING AQUACULTURE SYSTEM (RAS) FOR CRASSOSTREA VIRGINICA LARVAE

by

Isabelle Townsend

A Thesis
Submitted to the Graduate School,
the College of Arts and Sciences
and the School of Ocean Science and Engineering
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Science

Committee:

Dr. Leila Hamdan, Committee Chair
Dr. Kristina Mojica
Dr. Chet Rakocinski

August 2024

ABSTRACT

Recirculating Aquaculture Systems (RAS) offer control over physicochemical and biological factors impacting aquaculture success. One understudied facet of RASs is the water microbiome. Oyster larval microbiomes are shaped by their aquatic microbiomes. Understanding the dynamics of and the factors shaping RAS microbiome may provide insights to how the microbiome can be a tool for optimizing larval production. This study's goals were to investigate the microbiome stability of a RAS rearing oyster larvae across time and space, determine impacts of larval presence on the microbiome, and examine connections between larvae production outcomes and microbiome stability. Water samples were collected weekly from four compartments in the Thad Cochran Marine Aquaculture Center's RAS at the University of Southern Mississippi during three runs spanning May-October 2019 (168 days), June-October 2020 (115 days), and May-October 2021 (165 days). Sequencing was done using the V6-V8 variable regions of the 16S rRNA gene targeting bacteria. Water quality parameters and larval survival-assessments were collected. The RAS achieved stable physiocochemical conditions and microbiomes by the third run. Each compartment, with a specific function in water-reclamation/larval rearing, contained distinct microbiomes. Larval presence was not a factor shaping any compartment's microbiome. Connections between larvae and the Raceway (first water-reclamation stage) microbiome elevated over time, and instability loosely correlated with declining larval performance. Establishing baseline microbiomes present in closed-systems and further exploring connections between production and RAS microbiomes would enhance knowledge of RAS microbiome dynamics. The better the RAS microbiome is understood, the more valuable a tool in aquaculture it can be.

ACKNOWLEDGMENTS

I would like to acknowledge the mentorship of my Committee Members, Dr. Leila Hamdan, Dr. Kristina Mojica, and Dr. Chet Rakocinski. Their persistent support and guidance contributed to my knowledge and abilities as a scientist. I would like to acknowledge my labmates, Rachel Moseley, Dr. Rachel Mugge, and Kara Davis, for their daily support, mentorship, and assistance in completing this body of work. I would also like to thank the staff of the Thad Cochran Marine Aquaculture Center (TCMAC) for their work in sample collection and RAS maintenance throughout 2019, 2020, and 2021, especially Megan Gima and Benjamin Dosher who worked directly with me during this project. Thank you to NOAA for Graduate Assistantship support.

DEDICATION

This Masters Thesis is dedicated, firstly, to my grandfather who passed in the beginning of my journey as a graduate student. It is to Gray who I owe my admiration for the natural world and my passion for sailing. I was only able to tell you of one research cruise in the Gulf, but know I've carried you with me whenever I've sailed these past three years. I think of you every time I gaze at the sea, every time I lay under a blue-bird sky, and every time I impress colleagues with my knot-tying skills.

This work is secondly dedicated to the friends and family who emotionally supported me through this turbulent time of my life. I would not have seen the end of this journey as a graduate student without my parents, sister, and closest friends being a phone call or a drive away.

Lastly, I would like to thank my kitten Bodhi Gray for his insightful additions to my writing, such as ljnf;oansdfobdwf. He insists he deserves second authorship but is denied as he refused to contribute to lab work or data analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

AH	Algal Holding Tank
LT	Larval Tanks
RAS	Recirculating Aquaculture System
RW	Raceway
T	Tower
TCMAC	Thad Cochran Marine Aquaculture Center
USM	The University of Southern Mississippi

CHAPTER I – AQUACULTURE, OYSTER LARVAE, AND THE RAS MICROBIOME; AN INTRODUCTION

1.1 Overview of Aquaculture Systems

Aquaculture serves as vital means of obtaining seafood, and over time, several types of aquaculture systems have been developed ranging from little to full control over seafood production processes. Traditional open aquaculture systems typically employ the ambient environment (Tidwell, 2012a, Tidwell, 2012b) to maintain water parameters, nutrients, and waste removal. Therefore, these open systems are subject to variability from, for example, seasonality or extreme weather events (Tidwell, 2012b). A way of reducing variable conditions in aquaculture is to employ closed or semi-closed designs.

Ponds and flow-through systems, two types of semi-closed systems, provide methods to increase control over the production process (Fornshell *et al.*, 2012, Tucker & Hargreaves, 2012). A system is considered "semi" closed when reliance on ambient conditions is reduced via operator control over feed type and rates, water source, and water quality monitoring. These features combine to make systems less subjected to environmental variability, as observed in open systems. Ponds, for instance, are confined bodies of water with little input or outlet for the standing water, including rainfall and groundwater (Tucker & Hargreaves, 2012). The minimal influx of new water and the control of biotic factors can provide stable conditions for aquaculture, though these conditions cannot be controlled by the operator. Flow-through (FT) systems offer more direct control over water quality as they allow the operator to clean the water flowing into the rearing tank and discard the water after use (Fornshell *et al.*, 2012). Unlike ponds, FT systems have a continuous flow of water and have a higher production output, though

they are still reliant on the quality of the incoming water. Thus, they are subjected to the environmental conditions of the source water, which can lead to a crash (Gray *et al.*, 2022).

Recirculating aquaculture systems (RAS) are a closed-system approach to seafood production. Closed systems mitigate reliance on the ambient environment (Martins *et al.*, 2010, Ebeling & Timmons, 2012, Ahmed & Turchini, 2021). RASs can operate entirely indoors with the same water supply recycled through various water treatments. While both open and semi-closed systems offer varying degrees of control over production, RASs allow operators to hold the system conditions constant year-round without the effects of fluctuations in the ambient environment compared to open or semi-closed systems (Ebeling & Timmons, 2012, Tidwell, 2012b, Ahmed & Turchini, 2021). This ensures seafood being produced is maintained in stable conditions, with the long-term goal of mitigation production losses due to uncontrollable environmental events like storms (Ebeling & Timmons, 2012). Research into ideal conditions for various organisms in RASs has led to a better understanding of and, thus, a rise in RAS use for seafood such as Atlantic salmon, rainbow trout, and sturgeon in Nordic countries (Dalsgaard *et al.*, 2013).

Though RASs offer increased control over aquaculture production, RAS operations may be subject to mismanagement of system parameters, leading to poor water quality. Specific aspects of system designs, such as hydraulic pressure or physicochemical conditions, also vary depending on the organism being reared (Badiola *et al.*, 2012). High operation costs also are a barrier to RAS adoption (Badiola *et al.*, 2012) given that these systems have higher energy and maintenance demands than

traditional open or semi-closed systems (Midilli *et al.*, 2012, Tidwell, 2012b, Dalsgaard *et al.*, 2013, Ahmed & Turchini, 2021). The benefits of RAS may be promising from a seafood production perspective, but these benefits may not outweigh the costs of to construct and/or maintain a RAS (Badiola *et al.*, 2012, Dalsgaard *et al.*, 2013, Ahmed & Turchini, 2021).

A RAS involves a cyclic flow of water through successive water treatments prior to filling rearing tanks. A series of tanks are often used to reclaim used water to be treated and reused in the system (Ebeling & Timmons, 2012, Terjesen *et al.*, 2013, Balami, 2019). FT systems can have an initial control over the incoming water, such as disease treatments, particulate solid removal, and aeration (Fornshell *et al.*, 2012). However, the cyclic design of a RAS affords consistent and stable water conditions since the same water is being used and recycled during operation (Ebeling & Timmons, 2012). Additionally, RASs require fewer replenishments of the intake water compared to FT systems.

The conditions and methods required to rear specific organisms will depend on its specific needs. Commonly, mechanical filters remove suspended solids, biofilters are used to house nitrifying bacteria, which oxidize ammonia to nitrate, and heating/cooling systems are used to maintain the required temperature of the reared organism (Dalsgaard *et al.*, 2013). To prevent disease outbreaks, probiotics target specific pathogens known to harm the reared organism (Prado *et al.*, 2010, Stevick *et al.*, 2019). Ultraviolet (UV) treatment for pathogen removal can be effective at high doses, accomplishing its intended use in reducing overall microbial density (Laroche *et al.*, 2018), but has also been shown to reduce lobster larval survival when employed (Attramadal *et al.*, 2021). Designing a

RAS with a short hydraulic retention time (the amount of time water spends in each RAS compartment), however, has been shown to reduce the negative effects of UV on the microbial community and increase reared organism survival (Dahle *et al.*, 2022).

Aquaculture success is dependent on the successful production of larvae. RASs have been shown to yield significantly higher survival of cod (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014), lobster (Attramadal *et al.*, 2021), and lumpfish larvae (Dahle *et al.*, 2020) compared to FT systems. Attramadal *et al.* (2012) also noted, during the stressful transition to dry feed, cod larvae fared better in the RAS facility than the FT. All three studies demonstrate that a recirculating system benefits organism survival more so than the traditional FT system due to the more stable conditions achieved in closed system designs.

The success of aquaculture production is, in part, dependent on whether rearing conditions remain stable (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Bakke *et al.*, 2017). Production losses can occur when rearing conditions are inconsistent throughout production and/or when physicochemical conditions are unfavorable for the organism being reared (Terjesen *et al.*, 2013). Since RASs increase control over rearing conditions, a closed system approach to aquaculture may mitigate production losses. Commercial seafood production using RASs has been shown to be possible (Terjesen *et al.*, 2013), but for this to be feasible, water quality parameters need to be properly defined for the organism being reared (Dalsgaard *et al.*, 2013). In addition to water quality, the RAS design needs to favor the feeding rate, hydraulic retention, and other system parameters best suited for the organism being reared. For example, the optimal hydraulic retention time for salmon fry (Dahle *et al.*, 2022) is not the same for suspended oyster larvae (Qiu

et al., 2017). RASs have potential for widespread use in producing a variety of organisms, and this potential can be achieved via a better understanding of the specific RAS design and water quality parameters for each type of organism.

RASs have been used to successfully produce oysters, including *Crassostrea* species in several case studies. Kuhn *et al.* (2013) found growth and survival rates of adult *Crassostrea virginica* in a RAS were higher than previously published reports using open aquaculture systems. The advantage of using RAS to produce oysters year-round, even in winter when the low temperatures prevent open and semi-closed systems from successfully rearing oysters, was noted. In another study that compared a FT system (100% water renewal) to two RASs (25% and 0% water renewal), Asmani *et al.* (2016) found slightly higher success with RASs compared to FTs when rearing *Crassostrea gigas* larvae. In the two RASs, however, slower larval growth was observed. The non-optimal conditions cultivated in the closed systems, such as low pH and higher salinity, emphasize the need for more research on optimal rearing conditions for oyster larvae in RASs.

As with other species reared in RASs, there is a need to better understand the specific conditions required to optimally rear oysters. Ramos *et al.* (2021) showed rearing *C. gigas* larvae was feasible in RAS by defining the appropriate stocking density for this aquaculture method. While other larval culture systems support 1 to 10 larvae mL^{-1}, their RAS was able to successfully rear 50 to 75 larvae mL^{-1}. Having determined that stocking density and water quality were inversely related, a density of 50 larvae mL^{-1} was recommended to avoid water quality deterioration. Qiu *et al.* (2017) aimed to resolve issues relating to *Crassostrea angulata* larval production in RASs, namely by

determining optimal feeding rates and water flow. Since excessive feeding causes low water quality and low feeding rates cause lower larval growth, an optimal feeding rate was determined based on the ingestion rate of the larvae. Likewise, the flow rate for the RAS was determined based on larval swimming velocity. These aspects of the RAS can be decided specifically for the oyster species and oyster life stage being reared.

Though using RASs for oyster production has been deemed feasible, barriers remain to widespread use that require further research. Kuhn *et al.* (2013) reported toxic waste buildup in one study's recirculating system. This can be an issue for aquaculture facilities aiming to produce oysters for consumption (Chen *et al.*, 2016). Additional production issues, such as incorrect feeding rates, temperature, and water flow rates, can cause production failures in RASs, though these can be remedied once identified (Qiu *et al.*, 2017). Given the interest in production success with RAS, it would be valuable to define components and conditions of the system involved in successful system operation. Here, the primary component investigated was microbial community dynamics.

1.2 Overview of Oyster Microbiomes

Oysters are suspension feeders in both adult and larval stages and, as such, have constant contact with their surrounding environments, benthic or pelagic depending on their life stage (Pierce & Ward, 2018, Paillard *et al.*, 2022). Thus, both environmental (abiotic) conditions and biological factors influence the host microbiome of adult and larval oysters. To investigate microbial community dynamics in a closed system rearing oyster larva, knowing in what ways oyster larvae interact with their microbiota and the microbiome of their surrounding environment.

Oyster tissues, such as gills, mantle, inner shell, hemolymph, and gut, each contain specific microbiomes with some common taxa shared across the whole oyster (Hernández-Zárate & Olmos-Soto, 2006, King *et al.*, 2012, Lokmer & Wegner, 2015, Lokmer *et al.*, 2016a, Lokmer *et al.*, 2016b, Pierce *et al.*, 2016, Pimentel *et al.*, 2021). For example, the *C. virginica* gut compared to other tissues contains high relative abundances of Mollicutes and Chlamydiae, the mantle and gills contain a higher abundance of Spirochaetia, and the hemolymph contains Mollicutes, Chlamydiae, and Fusobacteriia (Pimentel *et al.*, 2021). Another study by King *et al.* (2012) supports the prevalence of Mollicutes in *C. virginica* stomach tissue and adds that Planctomycetes, Chloroflexi, Firmicutes, and Proteobacteria (Alpha-, Beta-, Delta-, and Gamma-Proteobacteria) were also observed.

The hemolymph microbiota interacts with the oyster host in response to stressors, such as location changes, temperature changes, and infection (Lokmer & Wegner, 2015, Lokmer *et al.*, 2016a, Lokmer *et al.*, 2016b, Green *et al.*, 2019). These stressors can cause not only mortality to the oyster but a lowering of alpha diversity in the hemolymph microbiome. Lokmer and Wegner (2015) found moribund oysters infected with a Vibrio strain harbored predominantly Arcobacters, a genus containing phylotypes of microaerophiles, and other genera containing opportunistic pathogens. Infection did not directly affect the hemolymph microbiome. Rather, oysters unable to survive the infection displayed a significantly lower microbial diversity than those who survived or were not infected (Lokmer & Wegner, 2015). This highlights a connection between the diversity of oyster tissue microbiomes and oyster performance that has been observed in

other animals (Chang *et al.*, 2008, Green & Barnes, 2010, Pillai *et al.*, 2014, Kanisan *et al.*, 2023).

The specificity of the stomach and gut tissue microbiomes potentially indicates a dependent relationship between host digestive tissues and the microorganisms inhabiting them (King *et al.*, 2012, Pierce *et al.*, 2016, Pimentel *et al.*, 2021). For example, Phaeobacter found in oyster gut tissues is hypothesized to provide defense against pathogenic Vibrio taxa (Pierce *et al.*, 2016, Pierce & Ward, 2019). High abundance of the class Mollicutes in the gut tissues indicates an association between nutrient acquisition mechanisms and Mollicutes (Pimentel *et al.*, 2021). Infection was associated with a lower abundance of Mollicutes in the gut and a decline in the overall health of the oyster, indicating the reliance oyster have on this gut tissue-microbiome relationship (Pimentel *et al.*, 2021).

The diet of an oyster has been shown to shape the digestive tissue microbiomes of adults (King *et al.*, 2012, Pierce *et al.*, 2016, Simons *et al.*, 2018) and the microbiome of larvae (Vignier *et al.*, 2021). For healthy, wild adult oysters, the gut microbiome consists of a rich community of Choroflexi, Planctomyctes, Spartobacteria, and Firmcutes among others (King *et al.*, 2012). The same study found the stomach microbiome to be more variable depending on location with one site represented predominantly by Mollicutes (>80%) and another by Planctomyctes (23-33%). This is in contrast to Pimental *et al.*(2021), where the focus on Mollicutes and the distinction between tissues revealed high relative abundances of Mollicutes and Chlamydiae, the latter of which King *et al.*(2012) found little of in either the stomach or gut. The high variability of these digestive tissues seen between different studies and sample sites is likely explained by the

variation in diet between locations. Changes made to the oyster diet can result in gut microbiome shifts even when all other conditions are held constant (Simons *et al.*, 2018) and this trend extends to the larval microbiome (Vignier *et al.*, 2021).

The oyster microbiome composition shifts (Asmani *et al.*, 2016) and declines in richness (Arfken *et al.*, 2021) as the larvae ages. Unlike settled adult oysters, larvae swim through the water column, and their interactions with the pelagic habitat provide a source for their larval microbiome (Asmani *et al.*, 2016, Laroche *et al.*, 2018).The period between larval hatching and settling is defined by three stages: D-stage (48 hours), Veliger (one week), and Pediveliger (two weeks). As the larvae progress through these stages, their microbiome lowers in diversity, becoming more specific to the individual due to a decline in bacterial colonization (Trabal Fernández *et al.*, 2014, Arfken *et al.*, 2021). Strong temporal changes in the larval microbiome indicate larval development is a driving force of change in their microbiome compositions (Stevick *et al.*, 2019, Arfken *et al.*, 2021).

Though the larvae and the water microbiomes are distinct (Asmani *et al.*, 2016, Stevick *et al.*, 2019, Arfken *et al.*, 2021), overlap between the two suggests the larvae may be selectively colonized by the bacteria in the water (Arfken *et al.*, 2021). The *C. virginica* larval core microbiome, as defined by Arfken *et al.* (2021) to be the OTUs found in 90% of samples, primarily consists of the families Rhodobacteraceae, Flavobacteriaceae, and Alteromonadaceae of the phyla Proteobacteria, Bacteroidota, and Pseudomonadota, respectively. This is consistent with Stevick *et al.* (2019) who found Proteobacteria dominated the larval bacterial communities. Both noted the larval microbiome consisted of a subset of the taxa present in the rearing water microbiome,

which consisted of the phyla Proteobacteria, Bacteroidota, Cyanobacteria, Actinobacteria, and Planctomycetes (Stevick *et al.*, 2019, Arfken *et al.*, 2021). Asmani *et al.* (2016) corroborates this, having found 60% of the *C. gigas* larval microbiome was accounted for in the seawater. It is important, then, when studying the cultivation of oyster larvae, to account for their life stage and the composition of the rearing water microbiome.

The oyster microbiome is indicative of the environment they were raised in (King *et al.*, 2012, Wegner *et al.*, 2013, Lokmer *et al.*, 2016a, Pierce *et al.*, 2016, Arfken *et al.*, 2021). Especially in the larval stages, the oyster microbiome is affected by the physiochemical conditions of their environment (Asmani *et al.*, 2016, Laroche *et al.*, 2018, Arfken *et al.*, 2021). In adulthood, their filter feeding activities allow the surrounding benthic environment to influence the composition of their pallial fluid and gut microbiomes (Pierce *et al.*, 2016). Relocation (Lokmer *et al.*, 2016b) and environmental condition changes (Wegner *et al.*, 2013, Lokmer & Wegner, 2015) cause shifts in the oyster microbiome composition. Even slight adjustments in an oyster's position in the intertidal zone significantly changes their digestive tissue microbiome (Offret *et al.*, 2020). This site-specificity extends to wild and aquaculture-raised oysters as individuals raised in aquaculture settings develop significantly different microbiomes than wild oysters (King *et al.*, 2012).

Closed-system approaches to aquaculture are designed to remove this environmental variability in seafood production. Thus, stressors, such as sudden temperature changes (Lokmer & Wegner, 2015) and seasonality (Pierce *et al.*, 2016) that lower the diversity of the microbiome composition are removed in RASs. While physiochemical conditions remain constant, the rearing water still contains a microbial

community which interacts with and influences the composition of larval microbiomes. Despite this, one aspect of closed aquaculture system research that is often overlooked is the water microbiome.

1.3 Overview of Aquaculture Microbiomes

Aquaculture microbiomes are shaped by the system design and water treatment methods used (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Asmani *et al.*, 2016, Almeida *et al.*, 2021) such that aquacultural practices and location have a significant impact on the composition of the water microbiome (Arfken *et al.*, 2021). RAS microbiomes are, in part, defined by the initial microbial community supplied to the system by the intake water (Bartelme *et al.*, 2019). Additionally, the RAS microbiome is sensitive to changes in water quality, such as changes in pH, temperature, and salinity (Blancheton *et al.*, 2013, Bakke *et al.*, 2017, Rud *et al.*, 2017, Mohamed Ramli *et al.*, 2018, Almeida *et al.*, 2021). High salinities, for example, cause lower microbial diversity in RASs (Bakke *et al.*, 2017, Rud *et al.*, 2017).

One important method of maintaining consistent water quality in aquaculture is UV sterilization, and its effect on RAS microbiomes is debated. UV can be used to reduce microbial density with varying outcomes on reared organism survival (Laroche *et al.*, 2018, Attramadal *et al.*, 2021, Dahle *et al.*, 2022). Attramadal *et al.* (2021) showed using UV within the water reclamation compartments resulted in a more variable microbiome compared to a RAS with the same intake water and no UV treatment. However, Dahle *et al.* (2022) saw no discernable effects of UV on the RAS microbiome aside from the intended reduction of total bacterial density.

RASs conserve water and system conditions; therefore, these systems may sustain a stable microbiome over time with stability defined as a lack of deviation from the baseline or norm. RASs greatly reduce the influx of water by retaining it, thus potentially allowing the microbiome to mature in the system over time (Attramadal *et al.*, 2012, Blancheton *et al.*, 2013, Attramadal *et al.*, 2014, Bakke *et al.*, 2017, Dahle *et al.*, 2022). Compared to FT systems where water flow is continuous, studies reveal some RASs support higher microbial diversity and, consequently, more stable communities over time (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Asmani *et al.*, 2017). Even when a FT was modified by adding a holding tank stage prior to the rearing tank (microbial maturation system), Attramadal *et al.* (2014) found microbial diversity and stability were higher in the RAS than the normal and modified FT systems. In that study, stability for each system was quantified by comparing the composition of a sample to another of the same system the following day, continuing until the last day of sampling. Termed a Moving Window Analysis on Bray-Curtis similarities, that study showed the RAS housed a significantly more stable microbiome over time compared to both FT systems. Further studies using Bray-Curtis similarities suggest RASs have a stable microbiome across multiple spawns (Dahle *et al.*, 2022) and across multiple semi-commercial systems (Bakke *et al.*, 2017). It is unknown whether stability connects to production outcomes in RASs, especially for oyster larval aquaculture. There is an apparent trend of higher oyster larval survival in RASs compared to FT systems (Asmani *et al.*, 2016), and oyster production in RASs has proven to be feasible (Qiu *et al.*, 2017, Ramos *et al.*, 2021). However, the microbial link between RAS microbiome stability and larval performance remains largely unknown.

CHAPTER II – CHARACTERIZING THE MICROBIOME OF A RAS ACROSS TIME AND SPACE TO UNDERSTAND RAS STABILITY OVER REPEATED RUNS

2.1 Characterizing the RAS Microbiome Across Time and Space

The stability of a RAS microbiome and the selective pressures of time and location in these closed systems are not well understood. Additionally, whether RAS microbiome stability connects with oyster larvae survival is not defined. The dynamics of a closed system may exhibit patterns of change associated with time, space, and larval production. By characterizing the RAS microbiome across time and space, insights into the stability of a RAS microbiome over time may be gained. It is hypothesized the RAS exerts selective pressure on the microbiome of water in the closed system across three periods of larval production (termed runs). Therefore, there will be no significant difference in the RAS microbiome by the end of each run across the system compartments. The null hypothesis then states the RAS microbiome is dictated by the introduction of source water supplied before the start of the first run and supplied as needed throughout the three runs. Therefore, there will be a significant difference in the RAS microbiome between the three runs across the system compartments.

It would be valuable to determine what features apply additional pressure on RAS microbial composition. This experiment will focus on the impact larval introductions have on the RAS microbiome. To investigate this, it is further hypothesized there will be a significant difference in the RAS microbiome in the presence or absence of larvae. The null hypothesis then states there will be no significant difference in the microbiome in the presence or absence of larvae.

2.2 Methods and Materials

2.2.1 RAS Design and Water Sampling

The RAS and water samples were provided by the University of Southern Mississippi's Thad Cochran Marine Aquaculture Center (TCMAC) located in Ocean Springs, MS. Water was collected weekly from multiple RAS compartments during larval production runs in May to October of 2019, June to October of 2020, and May to October of 2021 (Figure 2.1). Water in the RAS was continuously cycled through the within-loop compartments and was not disassembled between runs or brood spawns (Figure 2.1). Water moves from one compartment to the next on a 12-to-48-hour interval during the spawning season and once a week on the off season. Thus, water was continuously recycled between production runs and for the whole duration of the work with water introductions occurring as needed due to water loss.

Figure 2.1 *Timeline for water sampling*

Red indicates Run 1, orange indicates Run 2, and yellow indicates Run 3.

There are three stages of water treatment to reclaim water, denoted Raceway (RW), Tower 1 (T1), and Tower 2 (T2), as shown in Figure 2.2. Each stage had approximately the same volume of water per compartment with total water inventory

varying due to water loss and evaporation. Reclaimed water flowed into the Larval Tanks (LT), which was supplied algae fed from the Algal Holding Tank (AH). Used water from LT returned to RW to be reclaimed and reused.

New artificial seawater (ASW) was made in RW using salted well-water treated with Ethylenediaminetetraacetic acid (EDTA). Crystal Sea Bioassay Salt was added to the new water bring the salinity of the tank to between 17 and 19 ppt. EDTA was added to form chelates with metal ions to minimize trace metal contamination that would cause developmental issues with reared larvae (Utting & Helm, 1985). 1 mg of EDTA treats 1 L of new ASW.

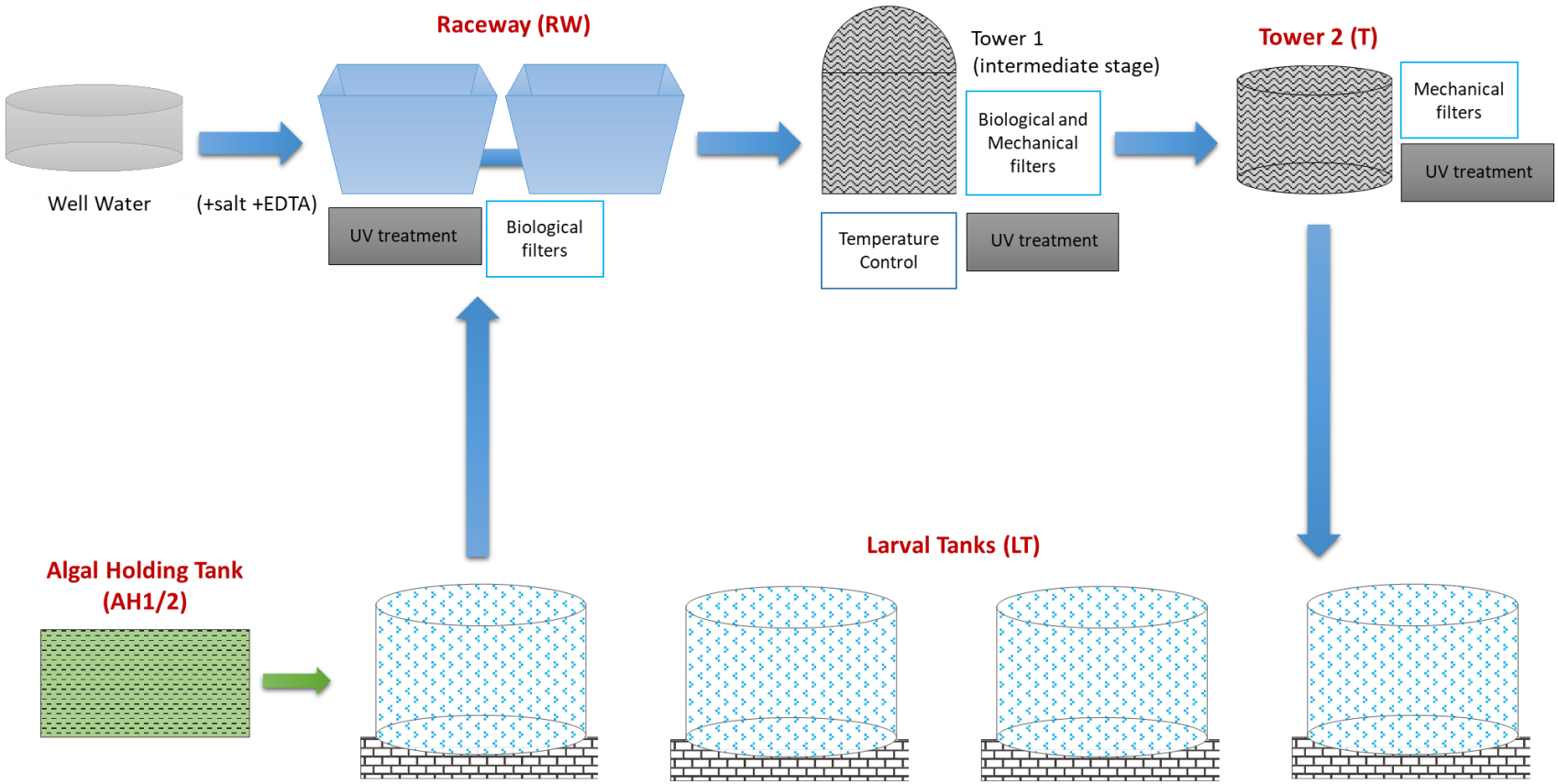

Figure 2.2 *Diagram of the RAS studied, including water flow and treatment per RAS compartment*

Blue arrows indicate the direction of water. Bolded red text indicates the compartments water samples were collected from for sequencing.

Each stage of water reclamation included filtration to clean and recondition the water for larval culture use in the LTs. The new and used ASW is exposed to UV sterilization and mechanical and biological filtration in RW. Tower 1 implemented redundant mechanical filtration, with bag filters, biological filtration via propeller washed bead filters (PBF-5), and UV treatment to further clean the water. Temperature was controlled by a Heat Flow and Physical Properties Package (HP3), maintained at 28 ℃. Tower 2, the final stage of reclamation, circulated water through a pleated canister filter, a canister filter filled with activated carbon, and a final canister filled with crushed coral. Temperature in T2 was also controlled by HP3 to maintain 28 ℃, and a redundant UV treatment further cleaned the water before it is pumped into the larval tanks.

Four 4400 L white polyethylene tanks, each with a fill volume of 4000 L, were used to rear oyster larvae except for in 2019, when due to lower larval counts, four smaller tanks were used for spawned broods. The reconditioned water from T2 was returned to the larval tanks (LT) while reclaimed effluent water from LT was pumped into the RW to begin the process of reconditioning the LT effluent to be recycled and reused. The number of tanks in use on a given day depended on the stocking density of the larvae. Tanks in use were drained every other day when larvae are under 150 microns and every day after larvae grow above 150 microns. During drain downs, the empty LT was rinsed with freshwater, disinfected with bleach, then rinsed again before being refilled by ASW from T2. Water and larvae alternated tanks following drain downs. One of the tanks was chosen at random for water sampling before drain downs. Algae was added to the tanks two to three times per day during larval production. Algae fed to reared larvae was harvested from batch or semicontinuous culture systems daily and

stored in 4 ℃ in the AH for a maximum of 24 hours. Unused algae in the AH were discarded.

2.2.2 Water Quality Parameters

The water quality parameters were tested daily in each compartment regardless of the sampling schedule. Temperature (℃), pH, salinity (ppt), and dissolved oxygen (mg/L) were assessed using a YSI Professional Plus handheld meter and Aquacheck strips. Ammonia (mg/L), nitrate (mg/L), and nitrite (mg/L) were assessed using a DR1900 Portable Spectrophotometer. Temperature controls aimed to maintain water temperature of 28 ℃.

2.2.3 Water Management

Water inventory primarily consisted of reconditioned, recycled water; however, when new water was needed and thus made, the date, time, and amount of salt added was recorded. Filters for RW, T1, and T2 were changed regularly, and the date and time of filter changes was recorded as well. Additional routine care, such as system sanitation involving rinsing lines and the floor with freshwater, was also recorded. Biosecurity is a high priority in RAS operations. Measures were taken to avoid or otherwise minimize contamination of the system, larvae, and algal feed.

2.2.4 Sample Collection and Handling

System water from RW, T2, and a randomly selected LT were collected for microbial sampling every Monday and Wednesday during runs. AH was sampled the day received (AH1) on Sunday/Tuesday and 24 hours after (AH2) on Monday/Wednesday. Six hundred mL of water from each of these compartments was collected and pumped through Millipore Sterivex filters (0.22 µm) (Millipore, Billerica, MA, USA) using a

peristaltic pump. Sterivex filters were stored at –80 ℃ until DNA is extracted and quantified. This totaled 455 samples across all compartments and runs.

2.2.5 DNA Extractions and Quantification

Cartridges containing filters were thawed at room temperature and cracked open using flame-sterilized pliers. The filters were removed from the cartridges using flame-sterilized scalpels and tweezers then placed in lysing matrix tubes for DNA extractions (Hamdan *et al.*, 2013). Genomic DNA was extracted from the filters using a modified version of the manufacturer's protocol of the BIO 101 FastDNA Spin Kit (MP Biomedicals, LLC, Santa Ana, CA, USA). Samples were processed on the FastPrep instrument at 4.5 rpm for 45 seconds to homogenize the sample and lyse the cells. This was then spun down using a centrifuge at 14000 RPM for 15 minutes, after which the supernatant of each sample was transferred to a clean tube. Protein Precipitation Solution (250 μL) was added to separate proteins from the DNA in the supernatant. Following a 5-minute centrifugation step (14000 rpm), the supernatant was poured into a clean tube and then combined with a binding matrix, to bind nucleic acids on solid media. After a period of rocking, settling, and supernatant removal, the sample was centrifuged for 1 minute, washed twice with concentrated SEWS-M washing solution, spun dry, then transferred to a catch tube to dry at room temperature for 5 minutes. DNAse free water (DES) (150 μL) preheated to 65°C was added to the dried sample then spun down for 1 minute. The resulting eluted DNA can then be stored at -20°C for short-term storage (days) or -80°C for long-term storage (years). The extracted DNA (ng/mL) of each sample was quantified using the manufacturer's protocol of the Qubit 2.0 Fluorometric Quantitation system (Invitrogen, Carlsbad, CA, USA). The quantity and purity of the DNA was assessed

using the NanoDrop spectrophotometer (ThermoFisher, Waltham, MA, USA). All samples were then plated on a 96-well plate (20 μL sample per well) for 16S rRNA gene sequencing and amplification.

2.2.6 16S rRNA Gene Sequencing of the V6-V8 Regions

Extracted DNA was then analyzed at the Integrated Microbiome Resource (IMR) at Dalhousie University in Nova Scotia, Canada. Specifically, the samples were subjected to amplification and sequencing of the 16S rRNA gene, according to the protocol from Comeau and company (2011). While the bacterial 16S rRNA gene contains nine variable regions, this study used the primers B969F (ACGCGHNRAACCTTACC) and BA1406R (ACGGGCRGTGWGTRCAA) to target the bacterial V6-V8 region. The regions typically targeted in oyster microbiome studies, primarily the V1-V2 and V4-V5 regions (Pierce *et al.*, 2016, Stevick *et al.*, 2019, Arfken *et al.*, 2021, Pimentel *et al.*, 2021), capture a wide varied of bacteria and archaea (Parada *et al.*, 2016, Walters *et al.*, 2016). The regions V3, V4, and V5 are also targeted in RAS microbiome studies (Attramadal *et al.*, 2014, Bakke *et al.*, 2017, Bartelme *et al.*, 2019, Dahle *et al.*, 2022). However, the present study used the V6-V8 regions covering specifically bacterial communities as the extraction protocol (Comeau *et al.*, 2011) is standard lab procedure, and these primers are able to distinguish a higher resolution of taxa in marine sediments (Parada *et al.*, 2016). Though the chosen regions are not standard for RASs or oysters, the less conserved regions offer a higher resolution of taxa, allowing more distinctions between lower taxonomic levels (Bukin *et al.*, 2019).

A total of 455 plated samples, as mentioned above, were processed via multiplexing enabling amplification of the 16S rRNA gene region (Meyer & Kircher,

2010). The PCR protocol includes an initial denaturation for a 30 s cycle at 98 ℃, followed by 30 cycles of denaturation at 98 ℃ for 10 s, a 20 s annealing cycle at 55 ℃, a 20 s elongation cycle at 72 ℃, and then a final extension for a 10 min cycle at 72 ℃ (Meyer & Kircher, 2010). PCR products were pooled to create one library for sequencing. The samples were loaded onto the Illumina MiSeq platform where 300 base pair paired-end sequences were produced for bioinformatic analysis.

2.2.7 Bioinformatics Analyses

Raw sequences were analyzed via a bioinformatics pipeline on the quantitative insights into microbial ecology 2 (QIIME2) version 2022.2 (Bolyen *et al.*, 2019) pipeline. The forward and reverse primers were removed with cutadapt (Martin, 2011) and reads lacking an adaptor were removed. The resulting trimmed reads were used to determine the parameters for DADA2 (Divisive Amplicon Denoising Algorithm 2). DADA2 were denoised, filtered, and merged the paired-end sequences to detect and correct Illumina amplicon sequence variant (ASV) data as well as filter chimeric sequences (Callahan *et al.*, 2016). Low quality regions of the sequences, determined by the trimmed reads data, were removed prior to the joining of paired reads. The output was a feature table, representative sequences, and denoising statistics. The feature table consisted of sequence frequencies per sample, which the representative sequences can be mapped to. Denoising statistics provided the percentage of reads filtered and correctly merged as the percent non-chimeric. These outputs were then be used to generate a phylogenetic tree using mafft, (Katoh & Standley, 2013) which was used to quantify diversity metrics via the core-metrics-phylogenetic plugin in QIIME2. A feature classifier was trained on the dataset and then used in reference to the SILVA database (version 132) at 99% similarity

based on the primer set used to target the V6-V8 regions (B969F/BA1406R) for taxonomic analysis (Quast *et al.*, 2012). The VSEARCH classifier (Rognes *et al.*, 2016) was used to assign taxonomy to the ASV feature table that DADA2 produced. Taxonomic classification was converted to relative frequency data (relative abundance) which can then be used for downstream analysis.

2.2.8 Data Quality Control

Following generation of the feature table from DADA2, sequence counts were analyzed to remove low sequence counts. Outliers were removed and the sampling depth was set based on the results of a median absolute deviation (MAD) method (Leys *et al.*, 2013). This was performed in R/RStudio to detect outliers using the median, rather than the mean, to determine the range of a dataset. The median absolute deviation and a researcher-determined coefficient are used to determine high- and low-count outliers. Leys (2013) outlines recommended coefficient values of 3, 2.5, and 2 as most to least conservative, respectively, but these recommendations are subjective depending on the dataset. A coefficient for each compartment was determined by weighing the cost of removing samples to the benefit of eliminating outliers, i.e. balancing the retention and refinement of data. Prospective coefficients retained 50% or more of data and set the lower limit (the sampling depth in QIIME2 core metrics) as high as possible to maximize sub-sampling. Sampling depth in the core metrics plugin is the frequency at which each sample is rarefied for subsequent diversity metric calculations, resulting in samples with sequence counts below the designated depth to be dropped. Selecting a low depth reduces the number of sequences within a sample that can be sub-sampled, retaining fewer

features than a higher depth. Thus, the low-count outliers detected by the MAD informed the sampling depths for each compartment.

2.2.9 Statistical Techniques

Using the QIIME2 core-metrics plugin, alpha diversity and beta diversity were calculated from the ASV feature table to understand diversity within and between samples, respectively. QIIME2 ASV data was used to construct a weighted unique fraction (UniFrac) distance matrix with the relative abundance data, which was then imported to Plymouth Routines in Multivariate Ecological Research 7 (PRIMER7) software (PRIMER-E Ltd., Plymouth, UK) for downstream analyses.

A BEST (or BIO-ENV) procedure was used to determine the best combination of given environmental variables that optimally match the microbial community structure being studied (Clarke & Ainsworth, 1993). Prior to the procedure being performed, environmental and microbial data were matched by sample date (day of sampling in a specific run) and compartment so that for every water quality measurement there was a corresponding water sample taken on the same day in the sample place. Thus, the Euclidean distance matrix constructed for environmental data and the UniFrac distance matrix of each compartment had corresponding data points. Spearman rank correlation coefficient were calculated for each environmental variable and all combinations of environmental variables. The output of the analysis is a model detailing the combinations of variables best fit to the biological data. This study aims to use this procedure to determine correlations between measured water quality parameters in the Raceway and Tower with the microbiomes of each respective compartment.

Bubble plots offer a visual means of visualizing differences in relative abundance of taxa between defined groups. Here, plots were generated in R/RStudio using average relative abundance in each compartment per run at the family level of taxonomy. The seventy-five most abundant families across all compartments in all runs were plotted.

2.2.10 RAS Microbiome Stability

To investigate trends in the microbiome of a RAS between runs and relative to compartment, data was statistically analyzed using the PRIMER7 software. The weighted UniFrac matrix imported into PRIMER7 was used to visualize the differences between samples with non-metric multi-dimensional scaling (nMDS). This was used to visually orient samples as two-dimensional and three-dimensional plots to aid in understanding how similar or different the microbial community composition is between samples, compartments, runs, and duration of runs.

Data was then statically analyzed using a permutational multivariate analysis of variance (PERMANOVA), which is a non-parametric test of differences in variance between variables within a factor(s). This statistical test does not assume normality and instead relies on permutations to analyze variance. The permutational/randomization approach assumes samples are independent and therefore exchangeable under a null of no difference by testing whether factor labels, determined by the experimenter, are exchangeable between samples (null accepted, $p>0.05$) or contribute to a factor's effect on the distribution of data (null rejected, $p<0.05$). Main tests were run using Type III (partial) sum of squares, fixed effects sum to zero for mixed terms, permutation of residuals under a reduced model, and 9999 permutations. Model factors included runs (year) and compartments (site) as fixed factors and days of runs as a random factor. To

elaborate, the “day of run” term consists of the day within one of the three runs a sample was taken, where Day 1 corresponds to the first day a sample was taken for each run. Pair-wise tests were then run on significant factors and interaction terms using the same parameters. Each factor and interaction term were assessed by considering the P-value and pseudo-F or t-test statistic.

A moving window analysis, informed by Attramadel *et al.* (2014), was used to assess changes in distances between samples as a means of investigating the stability of the RAS microbiome. The “window” is the pane through which a time series is viewed. To assess stability over time, the weighted UniFrac distance matrix from QIIME2 was subset by compartment then organized in chronological order. A custom script in R/RStudio selected distances between the first sample and the second, the second sample and the third, and so on until the end of the matrix was reached for every compartment. These successive distances were subset by each run (2019, 2020, 2021) and each “window” was displayed as a bar on a barplot. Additionally, the distances between the first sample taken in 2019 and all subsequent samples were selected and plotted to assess stability from the experiment’s starting point. Given UniFrac distances are defined as 0 being identical samples and 1 being completely different samples, high stability would be represented as short bars compared to less stable tall bars.

To statistically analyze the distances selected by the moving window analysis, t-tests were performed to determine the statistical differences between the means of each of the three runs. To statistically determine shifts in stability over time per compartment, successive distances were “smoothed” using Local Regression (LOESS), a non-parametric method that uses a least -squares regression calculation on local subsets of

data in a time series to smooth inconsistent or sporadic data. Following data smoothing, a Student's t-test Analysis of Regime Shifts (STARS) was performed to detect shifts in stability from the denoised time series. STARS is an algorithm built into the R/RStudio package rshift that detects changes in a time series, in this case UniFrac distances, and calculates a regime shift index (RSI) to quantify the magnitude of the shift(s). The output is the RSI of each time point where non-zero RSI values denote shifts. Here, this analysis was applied to successive UniFrac distances to detect changes in stability over time per compartment such that the output supplies specific "windows" in which a shift has occurred. This method of smoothing a time series to then apply regime shift detection was informed by Rodionov (2016).

2.2.11 K-Means Clustering

To further investigate the influence of factors driving data orientation, k-means clusters were generated for each compartment in PRIMER7. This algorithm assigns datapoints to specific *k* clusters by resemblance such that the clusters makeup samples with the lowest possible mean dissimilarity (distance from a local centroid) for *k* number of groups. This analysis visualizes the inherent partitions in a dataset, allowing comparisons to be made between algorithmically defined groups and levels of a tested factor. The number of clusters assigned to the algorithm was informed by a Silhouette Method performed in R. This method uses the minimum distance between samples and the average distance between each sample and all others to determine the optimal number of clusters for the dataset. The method was applied to each compartment, and a number of clusters to provide the algorithm was then determined based on the output.

2.2.12 Larval Presence and Absence

To determine connections between the RAS microbiome and larval presence, an analysis of similarity (ANOSIM) was used. This non-parametric test used the UniFrac matrix to rank dissimilarity between and among two groups and assess whether the similarity of samples between groups was greater than the similarity within groups. Data was subset by compartment. A one-way ANOSIM was run on the unordered factor present/absent with 9999 permutations.

2.3 Results

2.3.1 Data Quality Control

From the original 455 samples, samples labeled as weak and/or failed sequences were removed from the dataset, leaving 429 samples for analysis. The MAD approach was initially performed on the whole dataset. However, compartments exhibited varying sequence counts relative to each other (see Table 2.5) which skewed the calculation of the median. To account for these variations in sequence counts in each compartment, the approach was applied to each compartment separately (Table 2.1).

Multiple coefficients ranging from ± 0.5 to ± 3 (least to most conservative) were used to detect outliers. The coefficient was decided by finding a balance between outlier removal and sample retention. The decided coefficients were ± 0.9 for the RW and T, ± 1 for the LT, and ± 2 for the AH. This resulted in the retention of 73% of RW samples, 67% of T samples, 67% of LT samples, and 83% of AH samples for sample counts of 69, 64, 52, and 133, respectively. The lower limit established by the MAD then informed sampling depth in each compartment in the QIIME2 core metrics plugin. The results were

a sampling depth of 199 sequence counts in the RW, 833 sequence counts in the T, 9590 sequence counts in the LT, and 1553 sequence counts in the AH.

Table 2.1 *Results of the median absolute deviation (MAD) analysis performed on each RAS compartment to detect outliers*

A coefficient for the high and low ends was chosen according to the specific parameters set by each compartment.

MAD Results						
Compartment	MAD Coefficient	MAD High End	MAD Low End	% Retained	Sample Depth	Number of Samples
RW	± 0.9	16800.66	25.34	73%	199	69
T	± 0.9	10579.70	780.30	67%	833	64
LT	± 1.0	56811.37	9286.63	67%	9590	52
AH1 & AH2	± 2.0	63658.84	322.16	83%	1553	133
					Total All	318

A focus on sample retention in this study was important given the lack of replicate samples taken in each compartment at each sampling day. Any sample removed due to methodological error or outlier detection meant a time point in a specific sample site was lost. Removal of outliers using the MAD involved careful consideration of which samples were being removed and whether their designation as outliers was a feature of the RAS microbiome. Ultimately, the total 318 samples (Table 2.1) were decided upon. Removal of outliers, to reiterate, meant some time points were removed from each compartment. Therefore, the removal of outliers in each compartment resulted in inconsistent run lengths as described by Table 2.2. All subsequent downstream analyses use these 318 samples.

Table 2.2 *Run lengths in each compartment post-outlier removal*

Timeline per Compartment Post Outlier Removal					
Run	Original Run Length	Raceway	Tower	Larval Tank	Algal Holding Tank
2019	168 Days	Days 2 to 168	Days 2 to 168	Days 16 to 114	Days 8 to 168
2020	115 Days	Days 1 to 80	Days 29 to 115	Days 11 to 106	Days 2 to 108
2021	165 Days	Days 2 to 163	Days 2 to 165	Days 2 to 151	Days 1 to 163

2.3.2 Water Quality Parameters

Water quality parameters (temperature, salinity, pH, dissolved oxygen (DO), nitrate (NO_3), nitrite (NO_2), and ammonium (NH_3)) were monitored in the RW and T during all three runs. The average values for each run in both compartments are presented in Table 3. Temperature, overall, averaged 27.54 ± 2.03 ˚C with the highest annual average in 2020 and the lowest annual average in 2021. Temperature had the highest standard deviation in 2019 and lowest in 2021, suggesting 2019 temperatures were more varied and 2021 temperatures were more consistent. Salinity, overall, averaged 17.26 ± 1.56 ppt with the highest annual average in 2020 and the lowest annual average in 2021. Salinity had the highest standard deviation in 2019 and lowest in 2021, suggesting salinity was more varied in 2019 and more consistent in 2021. pH, overall, averaged 8.39 ± 0.15 with the highest annual average in 2020 and the lowest annual average in 2019. pH had relatively consistent standard deviations across all three runs. DO, overall, averaged 6.71 ± 0.62 mg/L with the highest annual average in 2021 and the lowest annual average in 2019. DO had the highest standard deviation in 2019 and lowest in 2020, suggesting DO was more varied in 2019 and more consistent in 2020.

Nitrate concentrations, overall, averaged 2.43 ± 1.87 mg/L with the highest annual average in 2019 and the lowest annual average in 2021. Nitrate had the highest

standard deviation in 2020 and lowest in 2021, suggesting nitrate concentrations were more varied in 2020 and more consistent in 2021. Nitrite concentrations, overall, averaged 0.03 ± 0.07 mg/L with the highest annual average in 2020 and the lowest annual average in 2021. Nitrite had relatively consistent standard deviations across all three runs. Ammonium concentrations, overall, averaged 0.80 ± 1.75 mg/L with the highest annual average in 2020 and the lowest annual average in 2021. Ammonium had the highest standard deviation in 2020 and lowest in 2021, suggesting ammonium concentrations were more varied in 2020 and more consistent in 2021. For all three parameters, standard deviations were lowest in 2021. This indicates that the last of the three runs had more consistent physicochemical conditions than the first two.

Table 2.3 *: Average water quality parameters per run*

Data for both the Raceway and Tower were included in these averages.

Average Physiochemical Conditions							
Runs	**Temp (˚C)**	**Salinity (ppt)**	**pH**	**DO (mg/L)**	**NO_3 (mg/L)**	**NO_2 (mg/L)**	**NH_3 (mg/L)**
2019	27.93 ± 2.19	17.18 ± 2.10	8.28 ± 0.16	6.55 ± 0.79	4.33 ± 1.30	0.03 ± 0.08	0.66 ± 1.29
2020	28.43 ± 1.90	18.20 ± 1.12	8.51 ± 0.11	6.60 ± 0.45	2.56 ± 2.04	0.05 ± 0.05	1.46 ± 2.31
2021	26.65 ± 1.59	16.70 ± 0.91	8.39 ± 0.11	6.91 ± 0.51	1.41 ± 0.98	0.01 ± 0.00	0.06 ± 0.35

Daily system conditions for temperature, salinity, pH, dissolved oxygen, nitrite, nitrate, and ammonium are presented in Figure 2.3 for the RW and T.

In the RW, temperature averaged 28.8 ± 1.8 ˚C in 2019, 29.5 ± 2.0 ˚C in 2020, and 26.5 ± 1.9 ˚C in 2021. In 2019, temperature rose from an initial 26.0 ˚C to a maximum of 32.9 ˚C by day 126, then fell to 22.9-26.8 ˚C by the end of the run. In 2020, temperatures began at a relatively consistent 29.4-31.7 ˚C until day 89 when temperatures fell from 28.9 ˚C (day 89) to a minimum of 23.6 ˚C (day 112). In 2021, temperatures were relatively low initially (23.2-26.7 ˚C) until rising after day 30. Between days 30 and

124, temperatures ranged from 26.9 ˚C to 29.4 ˚C. After day 124, temperature fell to 23.5-25.7 ˚C with a minimum of 20.5 ˚C (day 164).

In the T, temperature averaged 26.4 ± 1.9 ˚C in 2019, 27.3 ± 0.9 ˚C in 2020, and 26.8 ± 1.2 ˚C in 2021. In 2019, a maximum temperature occurred on day 152 when a temperature of 35.3 ˚C was measured, and a minimum occurred on day 164 when temperature of 15.6 ˚C was measured. In 2020, there was a slight rise in temperatures to 29.9-30.2 ˚C on days 19 and 20, and there was a fall to 24.4-27.3 ˚C between days 99 and 112. In 2021, a minimum temperature of 17.1 ˚C occurred on day 1, and a slight elevation to 27.5-28.5 ˚C occurred between days 93 and 105. Temperatures lowered in 2021 after day 125 to 22.6-26.6 ˚C.

In the RW, salinity averaged 17.40 ± 2.46 ppt in 2019, 18.32 ± 1.25 ppt in 2020, and 16.74 ± 0.89 ppt in 2021. In 2019, salinity began steadily rising from an initial 13.21-15.89 ppt to 17.25-20.17 ppt between days 30 and 140. Salinity rose after day 140, reaching a maximum of 27.65 ppt on day 153 and ending the run with a salinity of 25.08 ppt. In 2020, salinity rose from 16.48 ppt to 19.74 ppt between days 1 and 39. Salinity fell to a minimum of 12.67 ppt on day 57 before rising again to 19.74 ppt by day 103. From day 103 to the end of sampling, salinity remained at 17.61-18.59 ppt. In 2021, salinity ranged from 15.09 ppt to 18.02 ppt between days 1 and 123. With the exception of a local minimum of 13.61 ppt on day 147, salinity rose to 16.72-18.80 ppt by the end of the run.

In the T, salinity averaged 16.79 ± 1.18 ppt in 2019, 18.09 ± 0.96 ppt in 2020, and 16.67 ± 0.92 ppt in 2021. In 2019, salinity was initially low (14.40-15.97 ppt) and rose steadily to a maximum of 18.60 ppt by the end of sampling. In 2020, salinity was initially

high (18.07-19.41 ppt) until day 60 when salinity dropped to 15.75-16.97 ppt with a minimum of 15.16 ppt on day 80. After day 86, salinity began rising to 17.03-18.50 ppt. In 2021, salinity was initially 16.65 ppt until dropping to a minimum of 15.38 ppt by day 36. Salinity gradually rose over the course of the run to a maximum of 19.05 ppt by day 146.

In the RW, pH averaged 8.27 ± 0.15 in 2019, 8.50 ± 0.12 in 2020, and 8.37 ± 0.11 in 2021. In 2019, pH was at a minimum of 7.85 on day 25 and at a maximum of 8.55 on days 91 and 155-161. pH was generally 8.10-8.31 between days 1 and 126 then rose to 8.40-8.57 by the end of sampling. In 2020, pH rose from 8.24 to 8.59 between days 1 and 40 then fell to a minimum of 8.18 on day 57. pH rose again to a maximum of 9.49 on day 80 before falling again to 8.49 by the end of the run. In 2021, pH slightly lowered from an initial 8.46 to 8.06 between days 1 and 75 before rising again to 8.34-8.47, reaching a maximum of 8.75 on day 111. After day 131, pH lowered again to 8.06-8.43, reaching a minimum of 7.98 on days 146 and 148.

In the T, pH averaged 8.31 ± 0.17 in 2019, 8.52 ± 0.10 in 2020, and 8.41 ± 0.11 in 2021. In 2019, pH lowered from 8.26 to 8.01 between days 1 and 61. pH then gradually rose over the course of the run, reaching a maximum of 8.72 on day 164. In 2020, pH began at an annual minimum of 8.42 and did not rise higher than 8.60 for most of the run. A spike up to 9.49 occurred on day 74. In 2021, pH remained in the range of 8.31 to 8.53 for the majority of the run. A minimum of 7.84 occurred on day 74, and a maximum of 8.80 occurred on day 77.

In the RW, DO averaged 6.46 ± 0.61 mg/L in 2019, 6.43 ± 0.41 mg/L in 2020, and 6.94 ± 0.55 mg/L in 2021. In 2019, DO remained relatively consistent (6.40-6.99

mg/L) for the first 73 days. After day 73, DO began gradually lowering until day 149, reaching a minimum of 4.58 mg/L on day 142. DO rose for the remainder of the run to a maximum of 9.38 mg/L on day 167. In 2020, DO initially fell to 5.57-6.91 mg/L between days 1 and 39. DO then rose (6.26-6.61 mg/L between days 40 and 48) before falling again. Another period of rising DO concentrations took place between days 61 and 77 (6.52-7.81 mg/L) before lowering again to 5.89-6.96 mg/L between days 78 and 107. The run ended on a rise in DO of 7.08-7.65 mg/L. In 2021, DO remained mostly within the range of 6.87 mg/L to 8.42 mg/L until day 25. Between days 25 and 88, DO lowered to a range of mostly 6.31 mg/L to 6.90 mg/L with a local maximum of 8.79 on day 29 and a minimum of 3.65 mg/L measured on day 46. After day 88, DO gradually rose to 7.43 mg/L by the end of the run.

In the T, DO averaged 6.71 ± 1.01 mg/L in 2019, 6.77 ± 0.43 mg/L in 2020, and 6.87 ± 0.46 mg/L in 2021. In 2019, DO concentrations varied from lows of 4.57-5.84 mg/L and highs of 7.44-8.54 mg/L over the majority of the run. DO generally rose in the last 40 days of sampling to a range of 8.03-9.30 mg/L and reaching a maximum of 10.17 mg/L on day 164. In 2020, DO initially ranged from 6.03 mg/L to 7.31 mg/L in the first 70 days with a local drop to 4.46 mg/L on day 15. DO rose between days 71 and 75 to 7.36-7.80 mg/L then fell to mostly 5.98-6.58 mg/L between days 76 and 106. After day 106, DO rose to 7.29 mg/L by the end of sampling. In 2021, DO started at a maximum of 9.85 mg/L on day 1 then promptly fell to 6.92-7.24 mg/L between days 2 and 11. DO concentrations had sporadic lows of 4.87-5.63 mg/L occurring between days 17 and 74. Aside from these lows, DO remained mostly at 6.83-7.28 mg/L from day 12 to the end of sampling.

Nitrite concentrations in the RW averaged 0.06 ± 0.10 mg/L in 2019, 0.01 ± 0.01 mg/L in 2020, and 0.01 ± 0.00 mg/L in 2021. In 2019, nitrite concentrations initially rose from 0.06-0.07 mg/L to about 0.31 mg/L between days 30 and 70 before dropping to generally less than 0.01 mg/L for the remainder of the run. In 2020, nitrite concentrations remained below 0.03 mg/L (range of 0.01 mg/L to 0.03 mg/L) for the entire run. In 2021, nitrite concentrations also remained below 0.03 mg/L (range of 0.01 mg/L to 0.28 mg/L) for the entire run.

Nitrite concentrations in the T averaged 0.01 ± 0.01 mg/L in 2019, 0.08 ± 0.11 mg/L in 2020, and 0.01 ± 0.00 mg/L in 2021. In 2019, nitrite concentrations were initially high with a maximum of 0.08 mg/L on day 37 then fell to less than 0.02 mg/L (range of 0.00 mg/L to 0.02 mg/L) for the remainder of the run. In 2020, nitrite concentrations were initially below 0.02 mg/L between days 1 and 80. Between days 95 and 115, nitrate concentrations elevated to mostly 0.2-0.4 mg/L with a maximum of 0.45 mg/L on day 107. In 2021, nitrite concentrations generally lowered over the course of sampling from 0.02 mg/L to less than 0.01 mg/L.

In the RW, nitrate concentrations averaged 4.6 ± 1.4 mg/L in 2019, 1.4 ± 0.3 mg/L in 2020, and 1.5 ± 1.3 mg/L in 2021. In 2019, nitrate concentrations were initially low (2.0-2.2 mg/L) then rose to 4.8 mg/L between days 30 and 51. Nitrate continued to rise to a maximum of 7.1 mg/L between days 51 and 86 before falling to 3.0-4.6 mg/L over the rest of the run. In 2020, nitrate concentrations remained in a range of 1.1 mg/L to 1.3 mg/L for the first 60 days. For the rest of the run, nitrate rose to a range of 1.6 mg/L to 1.9 mg/L. In 2021, nitrate concentrations gradually fell from 1.4-1.8 mg/L to 0.5-0.9 mg/L between days 1 and 88. After day 88, nitrate rose to mostly 1.2-1.8 mg/L for the

remainder of the run. One sample taken on day 109 had a drastically higher concentration of nitrate (9.9 mg/L).

In the T, nitrate concentrations averaged 4.1 ± 1.2 mg/L in 2019, 3.2 ± 2.3 mg/L in 2020, and 1.3 ± 0.3 mg/L in 2021. In 2019, nitrate concentrations rose from 1.7-1.9 mg/L to a maximum of 6.3 mg/L between days 16 and 86. Nitrate then lowered over the remainder of sampling to 3.5-3.7 mg/L. In 2020, nitrate concentrations were initially 0.9-1.8 mg/L between days 1 and 80. Between days 95 and 115, nitrate concentrations elevated to about 4.8 mg/L with a maximum of 8.2 mg/L on days 98 and 99. In 2021, nitrate concentrations ranged from 0.6 mg/L to 1.9 mg/L throughout sampling. A slight dip (concentrations of 0.6-1.0 mg/L) occurred between days 30 and 100.

Ammonium concentrations in the RW averaged 0.7 ± 1.3 mg/L in 2019, 0.1 ± 0.3 mg/L in 2020, and 0.1 ± 0.4 mg/L in 2021. In 2019, ammonium was initially below 0.1 mg/L between days 30 and 72. Ammonium was elevated to 3.7-3.8 mg/L between days 77 and 91 then fell back to below 1 mg/L for the remainder of sampling. In 2020, ammonium concentrations varied sporadically from lows of 0-0.1 mg/L to highs of 0.2-0.3 mg/L throughout the run. A maximum of 0.6 mg/L occurred on days 59, 64, and 108. In 2021, ammonium concentrations were initially high at 0.5-1.7 mg/L between days 9 and 11 then fell to concentrations below 0.3 mg/L for the remainder of the run.

Ammonium concentrations in the T averaged 0.7 ± 1.3 mg/L in 2019, 2.2 ± 2.6 mg/L in 2020, and 0.0 ± 0.3 mg/L in 2021. In 2019, ammonium was initially below 1 mg/L between days 30 and 65. Ammonium was elevated to 3.1-3.8 mg/L between days 77 and 91 then fell back to below 0.1 mg/L for the remainder of sampling. In 2020, ammonium concentrations were initially below 0.5 mg/L between days 1 and 80.

Between days 95 and 115, nitrate concentrations elevated to mostly 3.0-7.1 mg/L. In 2021, ammonium was initially high at 1.4 mg/L on day 11 then lowered to below 0.3 mg/L for the remainder of sampling.

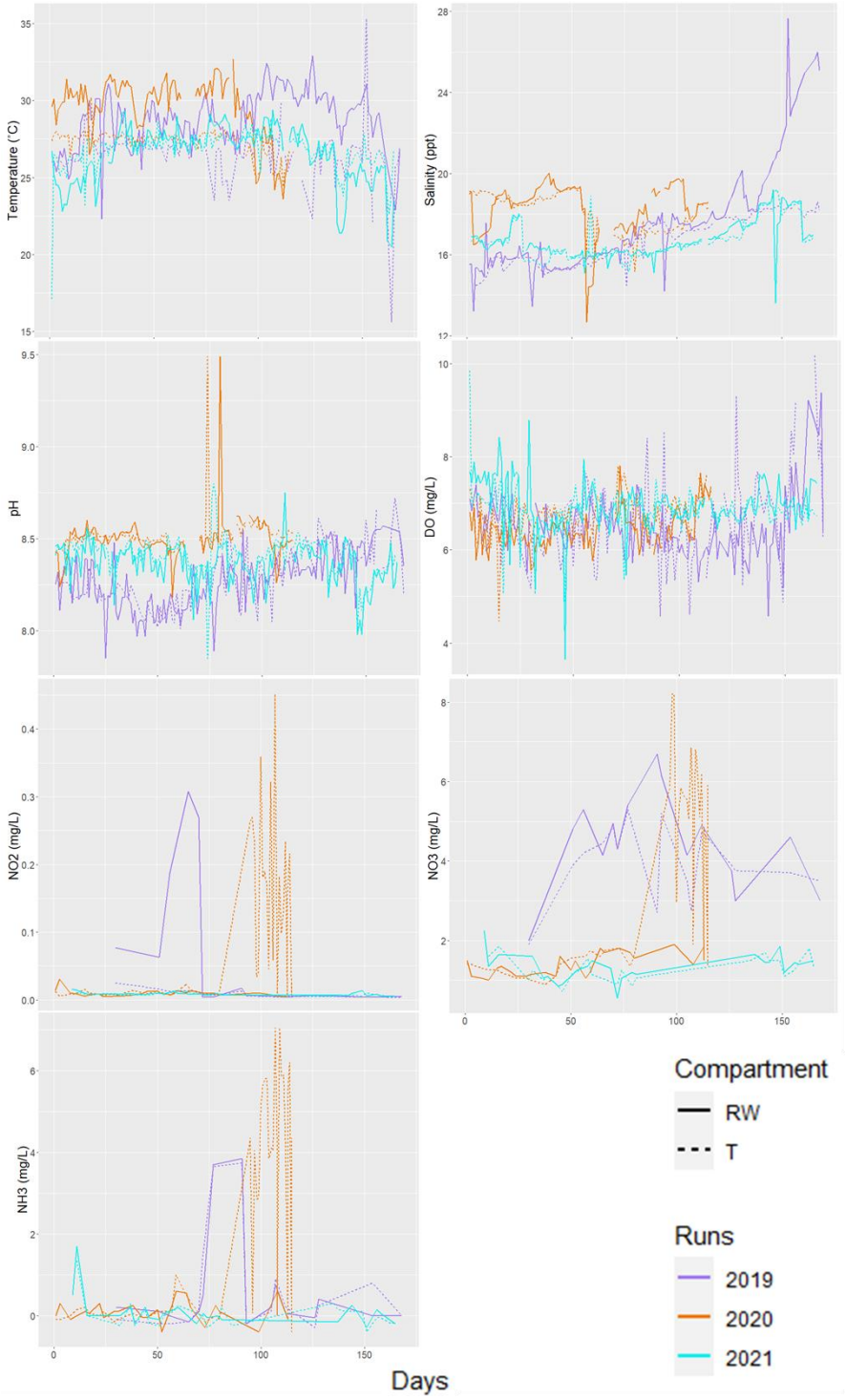

Figure 2.3 *Water quality parameters in the Raceway (RW) and Tower (T)*

Measured temperature (˚C), salinity (ppt), pH, dissolved oxygen (mg/L), nitrite (NO2) (mg/L), nitrate (NO3) (mg/L), and ammonium (NH3) (mg/L) as line plots per runs 2019 (purple), 2020 (orange), and 2021 (cyan) in the Raceway (RW) (solid lines) and Tower (T) (dashed lines). Physio-chemical parameters (y-axis) are plotted over each day in each of the three runs (x-axis).

2.3.3 Descriptive Statistics of the RAS Microbiome

Sequence count, Shannon diversity, observed ASVs, Chao1 index, and the ratio of observed ASVs to the Chao1 index for the whole system (all four compartments combined) for each run are reported in Table 2.4. The average overall sequence counts in samples from the whole microbiome dataset was 19,937.29 ± 3480.38 counts with the highest annual average in 2020 and lowest in 2019. Average overall Shannon diversity was 4.09 ± 0.29 with the highest annual average in 2019 and lowest in 2020. Average overall observed ASVs was 43.59 ± 3.49 with the highest annual average in 2019 and lowest in 2020. Average overall Chao1 Index was 124.90 ± 16.09 with the highest annual average in 2021 and lowest in 2019. This statistic was used to estimate how well the depth of coverage of the sequencing effort represented community composition and to determine the effectiveness of the sequencing method and primers. Chao1 index is calculated as a percentage of observed ASVs over Chao1. The average overall percent coverage for all compartments was 46.50 ± 5.14% with the highest annual average in 2019 and 2021. This indicates the sequencing effort underrepresented taxa present in samples, and the effectiveness of the chosen sequencing method was low. Assessing coverage per compartment would be valuable to determine whether particular components of the RAS microbiome are being over or underrepresented compared to other compartments.

Table 2.4 *Descriptive statistics for microbiome sequence data from the whole dataset*

Average Alpha Diversity Metrics in the Whole Dataset					
Runs	Sequence Count	Shannon Diversity	Observed ASV	Chao1 Index	Observed ASV/Chao1
2019	16108.95 ± 17830.04	4.43 ± 1.42	47.26 ± 27.82	110.49 ± 83.04	51.92 ± 20.76%
2020	22910.20 ± 17217.17	3.87 ± 1.58	40.32 ± 35.55	121.96 ± 121.84	45.86 ± 18.01%
2021	20792.72 ± 13925.54	3.98 ± 1.67	43.21 ± 36.03	142.26 ± 140.38	41.71 ± 17.84%

Table 2.5 reports average sequence count, Shannon Diversity, and Chao1 Index scores for the RW, T, LT, and AH individually for all three runs. In the RW, average overall sequence count was 6097.99 ± 2781.61 with the highest annual average in 2021 and lowest in 2019. Average overall Shannon diversity was 4.99 ± 0.34 with the highest annual average in 2020 and lowest in 2019. The average overall observed ASVs was 68.98 ± 98 with the highest annual average in 2020 and lowest in 2019. The average overall Chao1 score was 214.51 ± 80.77 with the highest annual average in 2021 and lowest in 2019. The average overall percent coverage was 39.90 ± 17.75% with the highest annual average in 2019 (60%) and lowest in 2021 (26%). This result suggests the sequencing effort underrepresented taxa in the RW microbiome.

In the T, average overall sequence count was 4805.58 ± 1722.34 with the highest annual average in 2020 and lowest in 2019. Average overall Shannon diversity was 6.00 ± 0.52 with the highest annual average in 2021 and lowest in 2020. Average overall observed ASVs was 82.81 ± 13.09 with the highest annual average in 2021 and lowest in 2019. The average overall Chao1 score was 235.62 ± 76.31 with the highest annual average in 2021 and lowest in 2019. The average overall percent coverage was 39.46 ± 11.47% with the highest annual average in 2019 (53%) and lowest in 2021 (34%). Like the RW, this result suggests the sequencing effort underrepresented taxa in the T microbiome.

In the LT, average overall sequence count was 32,641.08 ± 4667.95 with the highest annual average in 2019 and lowest in 2021. Average overall Shannon diversity was 4.14 ± 0.63 with the highest annual average in 2020 and lowest in 2021. Average overall observed ASVs was 40.95 ± 13.50 with the highest annual average in 2020 and lowest in 2021. The average overall Chao1 score was 167.85 ± 57.82 with the highest annual average in 2020 and lowest in 2021. The average overall percent coverage was 25.87 ± 2.96% with the highest annual average in 2021 (29%) and lowest in 2020 (24%). Like the RW and T, this result suggests the sequencing effort underrepresented taxa in the LT microbiome.

In the AH, average overall sequence count was 32061.37 ± 4232.81 with the highest annual average in 2019 and lowest in 2021. Average overall Shannon diversity was 2.80 ± 0.11 with the highest annual average in 2020 and lowest in 2021. Average overall observed ASVs was 15.29 ± 1.91 with the highest annual average in 2019 and lowest in 2020. The average overall Chao1 score was 29.95 ± 2.63 with the highest annual average in 2020 and lowest in 2019. The average overall percent coverage was 64.05 ± 13.96% with the highest annual average in 2019 (79%) and lowest in 2020 (51%). Percent coverage in all three runs ranged from lows of <30% and highs of >100%, indicating inconsistency in coverage from sample to sample. Though average coverage in the AH was about two times higher that of the RW, T, and LT, the sequencing effort was not consistently effective.

Table 2.5 *Descriptive statistics for microbiome sequence data from the RW, T, LT, and AH in each run*

Average Alpha Diversity Metrics of Each Run per Compartment						
Compartment	Runs	Sequence Count	Shannon Diversity	Observed ASV	Chao1 Index	Observed ASV/Chao1 (%)
Raceway	2019	3516.04 ± 4524.52	4.74 ± 0.96	55.96 ± 20.13	123.42 ± 92.12	60.02 ± 26.16
	2020	5734.42 ± 2382.58	5.37 ± 0.91	80.75 ± 25.03	242.72 ± 56.00	33.26 ± 7.80
	2021	9043.52 ± 4349.63	4.84 ± 1.14	70.24 ± 27.80	277.39 ± 110.03	26.43 ± 8.35
Tower	2019	2896.46 ± 2326.03	5.76 ± 0.76	72.71 ± 16.94	152.33 ± 63.84	52.63 ± 14.26
	2020	6242.75 ± 2030.21	5.65 ± 1.06	78.13 ± 17.77	252.36 ± 60.96	31.59 ± 7.67
	2021	5277.53 ± 2164.84	6.61 ± 0.41	97.59 ± 12.38	302.17 ± 75.67	34.16 ± 9.13
Larval Tanks	2019	36687.44 ± 10290.76	4.23 ± 1.05	39.78 ±16.86	165.28 ± 63.41	24.74 ± 5.79
	2020	33701.71 ± 21216.28	4.72 ± 1.73	55.00 ± 37.45	226.91 ± 134.63	23.63 ± 4.54
	2021	27534.08 ± 9887.45	3.46 ± 1.02	28.08 ± 16.99	111.35 ±111.88	29.23 ± 8.15
Algal Holding Tank	2019	36430.40 ± 16663.85	2.80 ± 0.87	17.16 ± 12.88	27.95 ± 20.57	78.65 ± 68.99
	2020	31774.33 ± 13321.76	2.90 ± 0.64	13.33 ± 3.74	32.92 ±29.85	50.83 ± 19.55
	2021	27979.39 ± 10333.54	2.68 ± 0.58	15.37 ±5.75	28.97 ± 21.84	62.66 ± 28.32

Shannon diversity in the RW, T, LT, and AH is plotted over time per run in Figure 4. In the RW, diversity in 2019 was generally variable, ranging from 2.74 to 6.00 for the first 91 days. The range of diversity narrowed (3.58 to 6.03) throughout the rest of the run. In 2020, diversity was initially sporadic, ranging from 3.26 to 6.28 in the first 24 days. For the rest of sampling, diversity ranged between 4.44 and 6.22. In 2021, diversity was initially high (6.28-6.58) for the first 16 days. From day 18 to 128, diversity generally lowered to, at minimum, 1.88 before rising to 3.89-5.29 by the end of sampling.

In the T, Shannon diversity in 2019 was initially sporadic, ranging from 3.63 to 6.61 for the first 23 days. Between days 37 and 58, diversity was a consistent 6.30-6.46 before falling to 4.79-5.30 between days 70 and 91. Diversity declined from 6.09 to 4.42 from day 100 to the end of sampling. In 2020, diversity ranged from 5.83-6.28 throughout the run with the exception of a sharp drop to 3.07 on day 108. In 2021,

diversity ranged from 6.04 to 7.17 throughout the run with the exception of a drop in diversity to 5.45 on day 58.

In the LT, Shannon diversity in 2019 rose from 2.48-2.85 to 4.22-5.60 over the course of sampling. In 2020, diversity was initially low (4.22) then rose to 6.43-6.51 between days 17 and 24. Diversity dropped to 2.39 on day 29 then rose slightly to 3.12-3.95 between days 45 and 106. In 2021, diversity mostly rose from about 2.89 to 5.61 between days 2 and 130. A spike in diversity of 7.40 occurred on day 100. Diversity lowered to 2.44-3.47 between days 135 and 144 before rising again to 3.84 by the end of sampling.

In the AH, Shannon diversity in 2019 ranged from lows of 1.81-2.35 to highs of 3.36-3.81 throughout sampling with the exception of a spike in diversity of 5.95 on day 65. In 2020, diversity ranged from lows of 1.45-2.50 and highs of 3.26-3.41 throughout the run. In 2021, diversity ranged from lows of 1.41-2.43 to highs of 3.08-4.35 and generally rose over the course of the run.

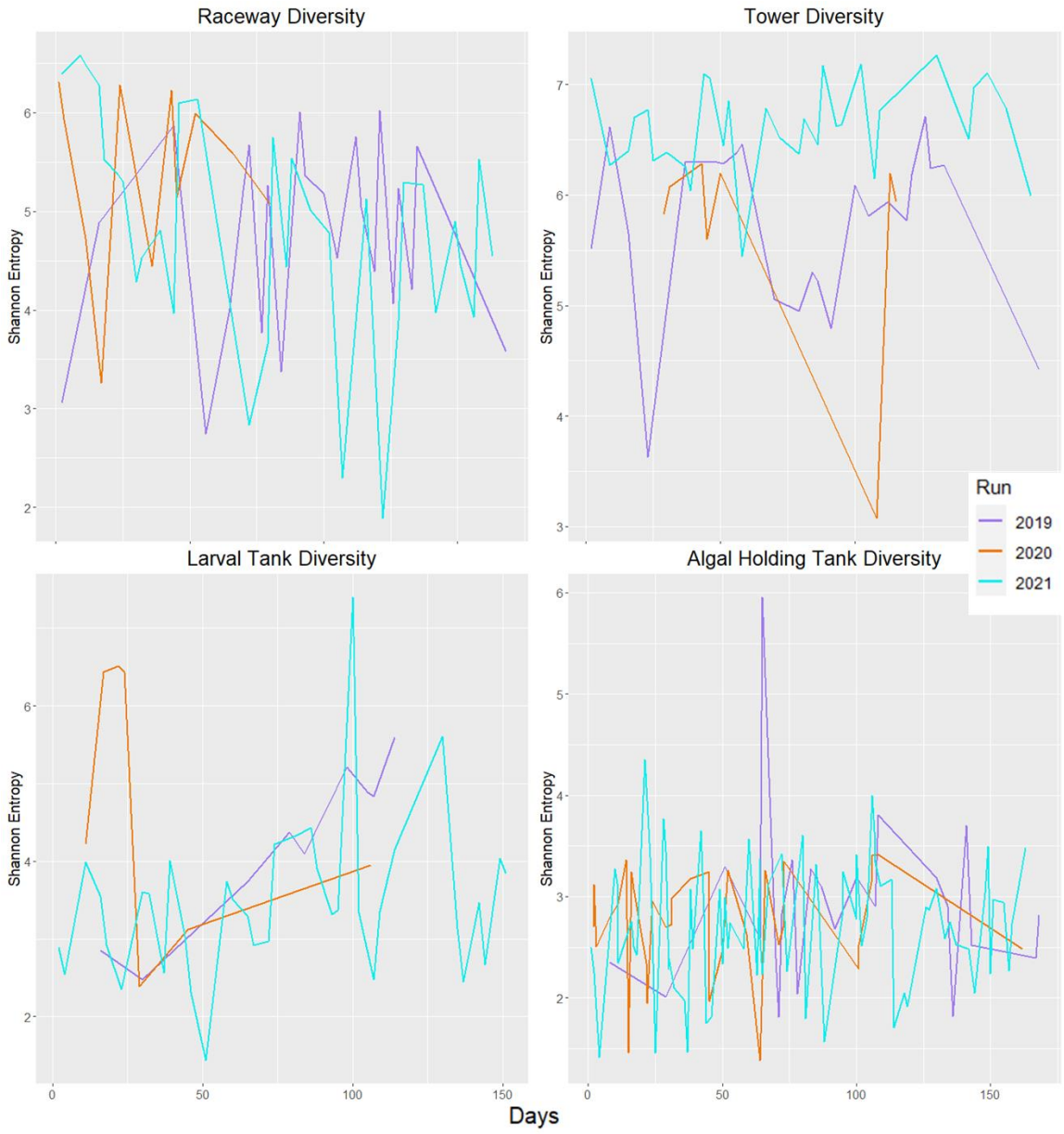

Figure 2.4 *Line plots depicting Shannon diversity per sample taken in each compartment*

Raceway (top-left), Tower (top-right), Larval Tanks (bottom-left), and Algal Holding tank (bottom-right) are plotted for 2019 (purple), 2020 (orange), and 2021 (cyan). Shannon diversity (y-axis) is plotted over the day each sample was taken (x-axis).

2.3.4 Community Composition in the RAS

Community composition was analyzed over time and space to determine similarities and differences in the RAS microbiome between runs and compartments.

Relative abundance and taxonomy data were displayed as the top 75 families present in each compartment for each run in Figure 2.5. This is to visualize community composition not to determine drivers of composition. Visually, Figure 2.5 shows little change in composition over time within compartments. Between compartments, RW and T have similar compositions while LT and especially AH have more sparce communities. Though there is some visible distinction between compartments, all four are dominated by classes Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia in all three runs, accounting for cumulatively 82.7% of the RW microbiome, 62.7% of the T microbiome, 95.8% of the LT microbiome, and 71.1% of the AH microbiome.

In the RW at the class level, Alphaproteobacteria consisted of 43.5% of the microbial community in 2019, 25.4% in 2020, and 23.4% in 2021. By contrast, Gammaproteobacteria made up 22.7% of the RW microbiome in 2019, 42.3% in 2020, and 41.6% in 2021. Bacteroidia consisted of 17.9% of the RW in 2019, 12.9% in 2020, and 18.3% in 2021. Deltaproteobacteria, Oxyphotobacteria, Actinobacteria, and OM190 consisted of between 6% and 1% of the RW in all three runs. The remaining classes had relative abundances of <1% individually. At the family level, the RW microbiome was comprised mostly of Rhodobacteraceae, Alteromonadaceae, Cellvibrionaceae, and Flavobacteriaceae. Rhodobacteraceae consisted of 33.3% of the microbial community in 2019, 9.2% in 2020, and 10.9% in 2021. Alteromonadaceae made up 5.9% of the community in 2019, 21.4% in 2020, and 5.3% in 2021. Cellvibrionaceae consisted of <1% of the community in 2019 and 2020 and 11.3% in 2021. Flavobacteriaceae comprised 9.2% of the community in 2019, 4.4% in 2020, and 4.4% in 2021. Several families, such as Spongiibacteraceae, Burkholderiaceae, Thalassospiraceae,

Xanthomonadaceae, and Pseudomonadaceae, had relative abundances of 1% to 6%. The remaining families had relative abundances of <1% individually.

In the T at the class level, Alphaproteobacteria consisted of 30.6% of the microbial community in 2019, 44.3% in 2020, and 23.2% in 2021. Gammaproteobacteria made up 26.3% of the T microbiome in 2019, 13.9% in 2020, and 15.3% in 2021. Bacteroidia consisted of 12.3% of the T in 2019, 12.0% in 2020, and 11.0% in 2021. Several classes made up between 10% and 1% of the T microbiome across all three runs, including Deltaproteobacteria, Oxyphotobacteria, Actinobacteria, Ignavibacteria, and Chlamydiae. The remaining classes comprise <1% individually. At the family level, the T microbiome was comprised mostly of Rhodobacteraceae, unclassified families in Oxyphotobacteria, Ignavibacteria, and Gammaproteobacteria, and an unclassified phylotype labeled "Other". Rhodobacteraceae consisted of 16.3% of the microbial community in 2019, 32.8% in 2020, and 6.9% in 2021. The Oxyphotobacteria family consisted of 4.6% of the community in 2019, 7.7% in 2020, and 7.8% in 2021. The Ignavibacteria family comprised of less than 1% of the community in 2019, 4.0% in 2020, and 7.3% in 2021. The Gammaproteobacteria family consisted of 4.9% of the community in 2019, 6.3% in 2020, and 5.6% in 2021. The "Other" category made up 10.0% of the community in 2019, 5.9% in 2020, and 11.2% in 2021. The remaining families had relative abundances of <5% individually.

In the LT at the class level, Alphaproteobacteria consisted of 50.0% of the microbial community in 2019 and 2020 and 49.2% in 2021. Gammaproteobacteria made up 30.2% of the LT microbiome in 2019, 33.4% in 2020, and 32.9% in 2021. Bacteroidia consisted of 17.0% of the LT in 2019, 12.0% in 2020, and 13.4% in 2021.

Oxyphotobacteria comprised an average 2.6% of the microbiome in all three runs, and the remaining classes comprised <1% individually. At the family level, the LT microbiome was comprised mostly of Rhodobacteraceae, Alteromonadaceae, Flavobacteriaceae, Thalassospiraceae, and Pseudomonadaceae. Rhodobacteraceae consisted of 49.9% of the microbial community in 2019, 22.8% in 2020, and 37.5% in 2021. Alteromonadaceae consisted of 15.9% of the microbial community in 2019, 22.7% in 2020, and 13.8% in 2021. Flavobacteriaceae consisted of 10.2% of the microbial community in 2019, 3.0% in 2020, and 7.9% in 2021. Thalassospiraceae consisted of 3.2% of the microbial community in 2019, 23.8% in 2020, and 8.1% in 2021. Pseudomonadaceae consisted of 1.7% of the microbial community in 2019, 5.0% in 2020, and 9.9% in 2021. The remaining families had relative abundances of <6% individually.

In the AH at the class level, Alphaproteobacteria consisted of 11.1% of the microbial community in 2019, 26.1% in 2020, and 16.0% in 2021. Gammaproteobacteria made up 23.1% of the AH microbiome in 2019, 19.7% in 2020, and 22.0% in 2021. Bacteroidia consisted of 39.3% of the T in 2019, 28.6% in 2020, and 27.3% in 2021. Oxyphotobacteria consisted of 26.3% in 2019, 25.6% in 2020, and 30.2% in 2021. Aside from Bacilli having an abundance of 2.74% in 2021 and Actinobacteria having an abundance of 1.33% in 2021, all other classes had an individual abundance of <1%. At the family level, the AH microbiome was comprised mostly of a family in Oxyphotobacteria, Alteromonadaceae, Flavobacteriaceae, Cyclobacteriaceae, Thalassospiraceae, and Rhodobacteraceae. The Oxyphotobacteria family consisted of 26.3% of the community in 2019, 25.6% in 2020, and 29.9% in 2021. Alteromonadaceae

consisted of 11.6% of the community in 2019, 11.5% in 2020, and 14.9% in 2021. Flavobacteriaceae consisted of 21.2% of the community in 2019, 5.7% in 2020, and 7.8% in 2021. Cyclobacteriaceae consisted of 9.8% of the community in 2019, 9.1% in 2020, and <0.5% in 2021. Thalassospiraceae consisted of 5.4% of the community in 2019, 22.5% in 2020, and 2.6% in 2021. Rhodobacteraceae consisted of 4.8% of the community in 2019, 2.3% in 2020, and 11.0% in 2021. The remaining families had relative abundances of <7% individually.

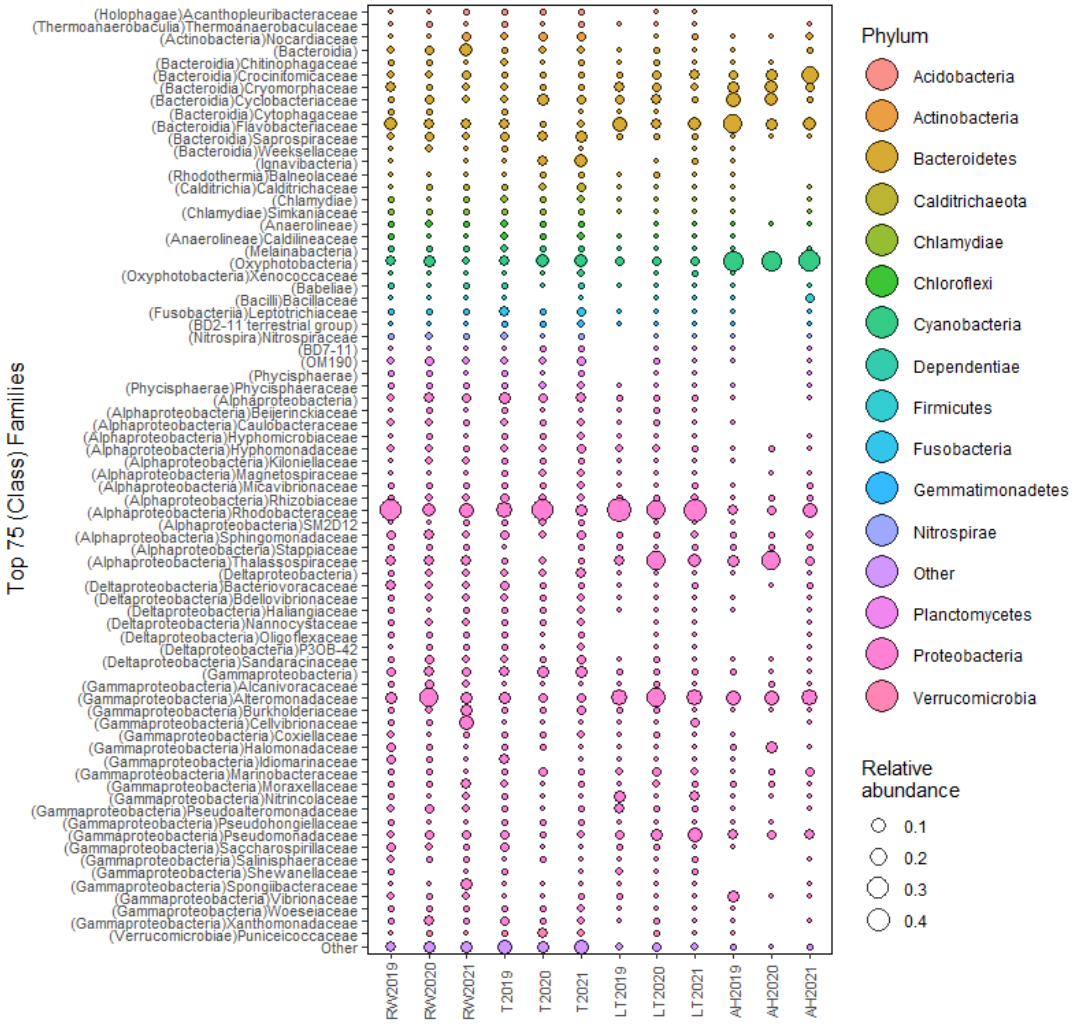

Figure 2.5 *Bubble Plot of relative abundance of the top 75 families (y-axis) for each compartment*

The top 75 families (y-axis) for each compartment (Raceway (RW), Tower (T), Larval Tanks (LT), and Algal Holding tank (AH), from left to right) per run (2019, 2020, and 2021, from left to right). Classes denoted in parentheses (y-axis). Bubbles colored by phylum. Average relative abundance is represented by bubble size.

The whole microbiome dataset, including all runs and compartments, was visualized using nMDS plots (Figure 2.6) to determine orientation of samples to each other. In these plots, samples are identified according to the compartment they were collected in (Figure 2.6, top) and the run collected in (Figure 2.6, bottom). The majority

of RW samples grouped with LT and T samples at a UniFrac distance of 0.45. Several RW samples were interspersed on the left side of the graph. The majority of samples from the T formed a cluster at a distance of 0.4, although there were several samples that did not adhere to this cluster and instead were interspersed throughout the left side of the graph. The majority of LT samples were spread amongst RW, T, and AH2 samples on the right side of the graph. AH1 samples were interspersed throughout the graph, mostly on the top and left side. AH2 samples mostly grouped together at a distance of 0.45 above RW and LT samples. Several AH2 samples were interspersed amongst AH1 samples on the left side of the graph.

When samples were analyzed for the effect of run year (Figure 2.6, bottom), there was no clear pattern in the distribution based on UniFrac distance. The same was observed for day of run, where no clear trend in ordination based on sampling day was observed across or within runs.

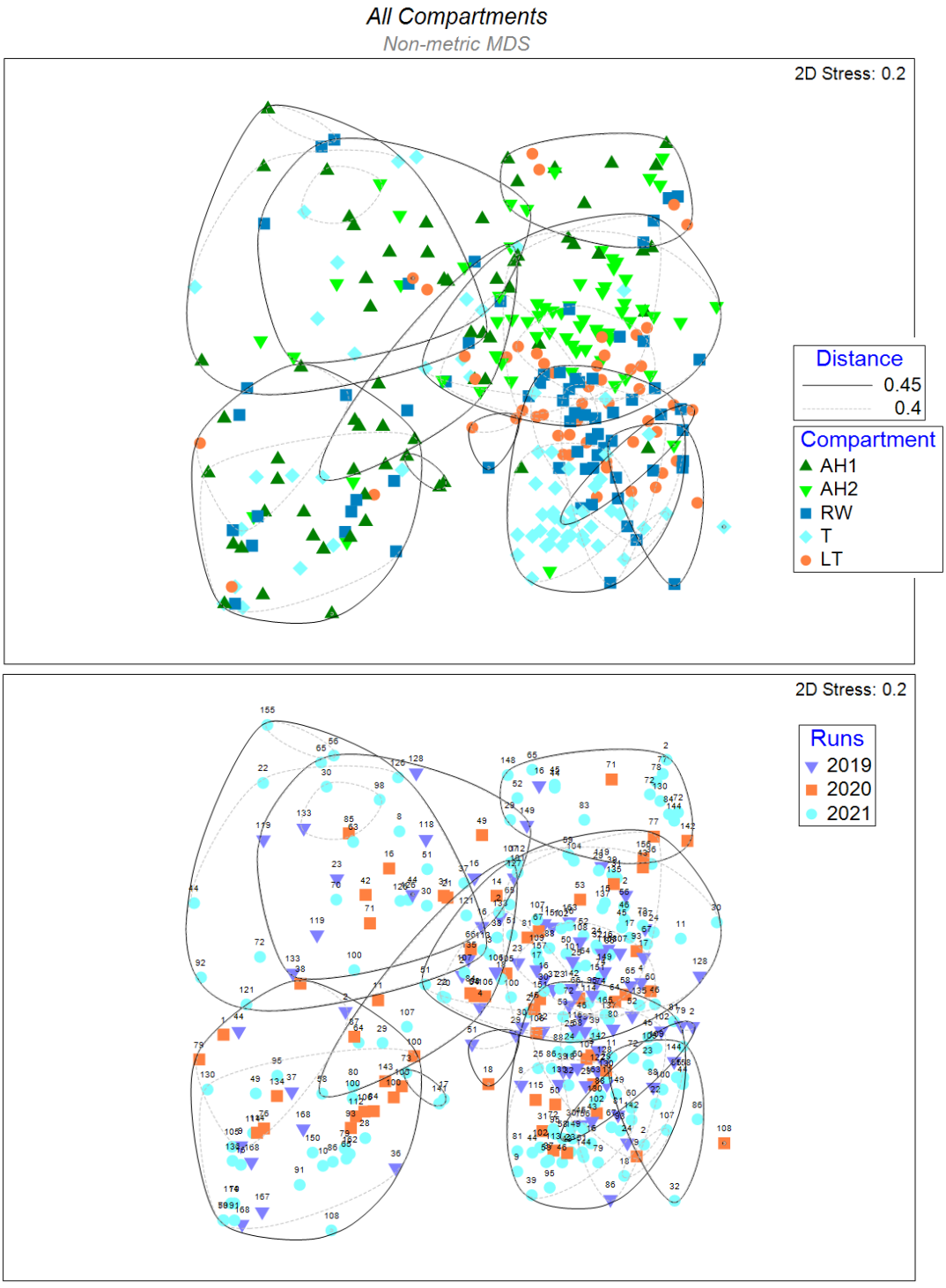

Figure 2.6 *Non-metric multidimensional scaling analysis (nMDS) for all compartments across all runs*

Distances between samples from the weighted UniFrac distance matrix. Samples are grouped by compartments (top) and runs (bottom) for visualization. Numbers above each symbol indicate the day the sample was taken within its run. A CLUSTER analysis generated distance contour lines at 0.45 (55% similar) and 0.4 (60% similarity).

The PERMANOVA main test on the data included in Figure 2.6 is presented in Table 2.6. The results of the analysis indicate the compartment and the run samples were collected from are significant drivers of community composition. The lack of a discernable pattern of orientation seen in the nMDS plots, especially orientation by run (Figure 2.6, bottom), may be attributed to their nearly equal contribution to community composition, as evidenced by the close Pseudo-F values for compartment (F=5.2836) and run (F=5.2164). The analysis confirms there is a significant interaction between the two terms.

Table 2.6 *Permutational multivariate analysis of variance (PERMANOVA) main test performed on the whole dataset to determine differences in the microbiome by runs and compartments*

Bolded text indicates a significant result (p<0.05).

Main Test PERMANOVA on the Whole Dataset						
Sums of squares type: Type III (partial)						
Fixed effects sum to zero for mixed terms						
Permutation method: Permutation of residuals under a reduced model						
Number of permutations: 9999						
Factors						
Name	Type	Levels				
Runs	Fixed	3				
Compartment	Fixed	5				
Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Runs	2	1.2656	0.63278	5.2164	**0.0001**	9908
Compartment	4	2.5637	0.64092	5.2836	**0.0001**	9882
Runs x Compartment	8	2.1976	0.27471	2.2646	**0.0001**	9854
Res	303	36.755	0.1213			
Total	317	44.945				

To differentiate the interaction effect between factors compartments and runs, PERMANOVA pairwise tests were performed for pairs of compartments (Table 2.7) and pairs of runs (Table 2.8). In 2019, results show eight out of the ten pairs were different,

the exceptions being AH1 and T and the AH1 and AH2. The t-test indicates the T and LT had the most distinct communities. In 2020, eight out of the ten pairs were different, the exceptions being the RW and LT and the LT and AH2. The t-test indicates the T and AH1 were the most differentiated. By 2021, all compartments were different from one another with the t-test indicating the T and AH2 were the most differentiated.

In three of the four compartments (RW, T, and LT), 2019 was different from 2020 and 2021. In the RW, 2020 and 2021 were not significantly different, and the higher t-test score for 2019 and 2021 indicates the first and last runs had the most distinct communities. In the T, all runs were significantly different. The t-test indicates 2019 and 2020 were the most distinct. In the LT, all runs were significantly different, and like the RW, the t-test indicates 2019 and 2021 are the most distinct. In AH1 and AH2, 2019 and 2021 was the only pair not significantly distinct from each other, and the t-tests for both indicate 2020 and 2021 had the most distinct communities.

Table 2.7 *Permutational multivariate analysis of variance (PERMANOVA) pair wise tests were performed on the Whole Dataset for factors runs and compartments within levels of runs*

Bolded text indicates a significant result ($p<0.05$).

Pair Wise Test PERMANOVA on the Whole Dataset: Levels of Compartment			
Sums of squares type: Type III (partial)			
Fixed effects sum to zero for mixed terms			
Permutation method: Permutation of residuals under a reduced model			
Number of permutations: 9999			
Factors			
Name	Type	Levels	
Runs	Fixed	3	
Compartment	Fixed	5	
Term 'Runs x Compartment' for pairs of levels of factor 'Compartment'			
Within level '2019' of factor 'Runs'			
Groups	t	P(perm)	Unique perms
AH1, RW	1.3835	**0.045**	9905
AH1, T	1.1801	0.2039	4322
AH1, LT	1.5845	**0.0073**	9692
AH1, AH2	1.3626	0.0589	9076
RW, T	2.5464	**0.0004**	9928
RW, LT	1.4378	**0.014**	9922
RW, AH2	1.8003	**0.0017**	9915
T, LT	2.8784	**0.0001**	9919
T, AH2	2.4965	**0.0003**	9914
LT, AH2	1.3152	**0.0456**	9900
Within level '2020' of factor 'Runs'			
Groups	t	P(perm)	Unique perms
AH1, RW	1.6838	**0.0181**	9943
AH1, T	2.6645	**0.0001**	9927
AH1, LT	1.4464	**0.0624**	9931
AH1, AH2	2.1232	**0.001**	9918
RW, T	1.7225	**0.0078**	9836
RW, LT	0.58696	0.9727	9794
RW, AH2	1.6058	**0.0235**	9004
T, LT	1.5675	**0.0154**	9397
T, AH2	2.4012	**0.0004**	7776
LT, AH2	1.3442	0.0782	6633
Within level '2021' of factor 'Runs'			
Groups	t	P(perm)	Unique perms
AH1, RW	1.5859	**0.0353**	9931
AH1, T	3.2782	**0.0001**	9942
AH1, LT	2.8876	**0.0001**	9933
AH1, AH2	2.2731	**0.0005**	9952
RW, T	2.3349	**0.0001**	9919
RW, LT	1.7492	**0.0028**	9937
RW, AH2	1.9745	**0.0008**	9941
T, LT	2.5228	**0.0001**	9932
T, AH2	3.925	**0.0001**	9933
LT, AH2	2.7062	**0.0001**	9929

Table 2.8 *Permutational multivariate analysis of variance (PERMANOVA) pair wise tests were performed on the Whole Dataset for factors runs and compartments within levels of compartment*

Bolded text indicates a significant result ($p<0.05$).

Pair Wise Tests PERMANOVA on the Whole Dataset: Levels of Runs			
Sums of squares type: Type III (partial)			
Fixed effects sum to zero for mixed terms			
Permutation method: Permutation of residuals under a reduced model			
Number of permutations: 9999			
Factors			
Name	Type	Levels	
Runs	Fixed	3	
Compartment	Fixed	5	
Term 'Runs x Compartment' for pairs of levels of factor 'Runs'			
Within level 'RW' of factor 'Compartment'			
Groups	t	P(perm)	Unique perms
2019, 2020	1.41	**0.0411**	9911
2019, 2021	1.8833	**0.0012**	9919
2020, 2021	1.2225	0.1541	9942
Within level 'T' of factor 'Compartment'			
Groups	t	P(perm)	Unique perms
2019, 2020	2.7897	**0.0001**	9439
2019, 2021	2.6568	**0.0003**	9948
2020, 2021	1.7546	**0.0076**	9905
Within level 'LT' of factor 'Compartment'			
Groups	t	P(perm)	Unique perms
2019, 2020	1.5179	**0.0034**	9912
2019, 2021	1.98	**0.0001**	9942
2020, 2021	1.5813	**0.0105**	9927
Within level 'AH1' of factor 'Compartment'			
Groups	t	P(perm)	Unique perms
2019, 2020	1.3782	**0.0894**	9834
2019, 2021	1.0888	0.2874	9947
2020, 2021	1.6836	**0.0276**	9951
Within level 'AH2' of factor 'Compartment'			
Groups	t	P(perm)	Unique perms
2019, 2020	1.4266	**0.0289**	9631
2019, 2021	1.2102	0.1544	9929
2020, 2021	1.5798	**0.0205**	9932

2.3.5 RW Microbiome

The RW dataset was visualized via an nMDS plot (Figure 2.7) to determine orientation of samples to each other based on time. This was visualized by coloring

samples according to the run they were collected in and labeled by the day of sampling. Samples from 2019 are spread across the lower half of the plot in three clusters at a distance of 0.4 (60% similarity). 2020 samples do not form a distinct group, though three samples form an outgroup at a distance of 0.4 with a 2021 sample. Generally, 2020 samples orient between 2019 and 2021 samples. Samples from 2021 form five tight clusters in the top half of the plot at a distance of 0.3 (70% similarity) and one out group at the top of the graph. There is a clear distinction between samples taken in 2019 and 2021 in the RW.

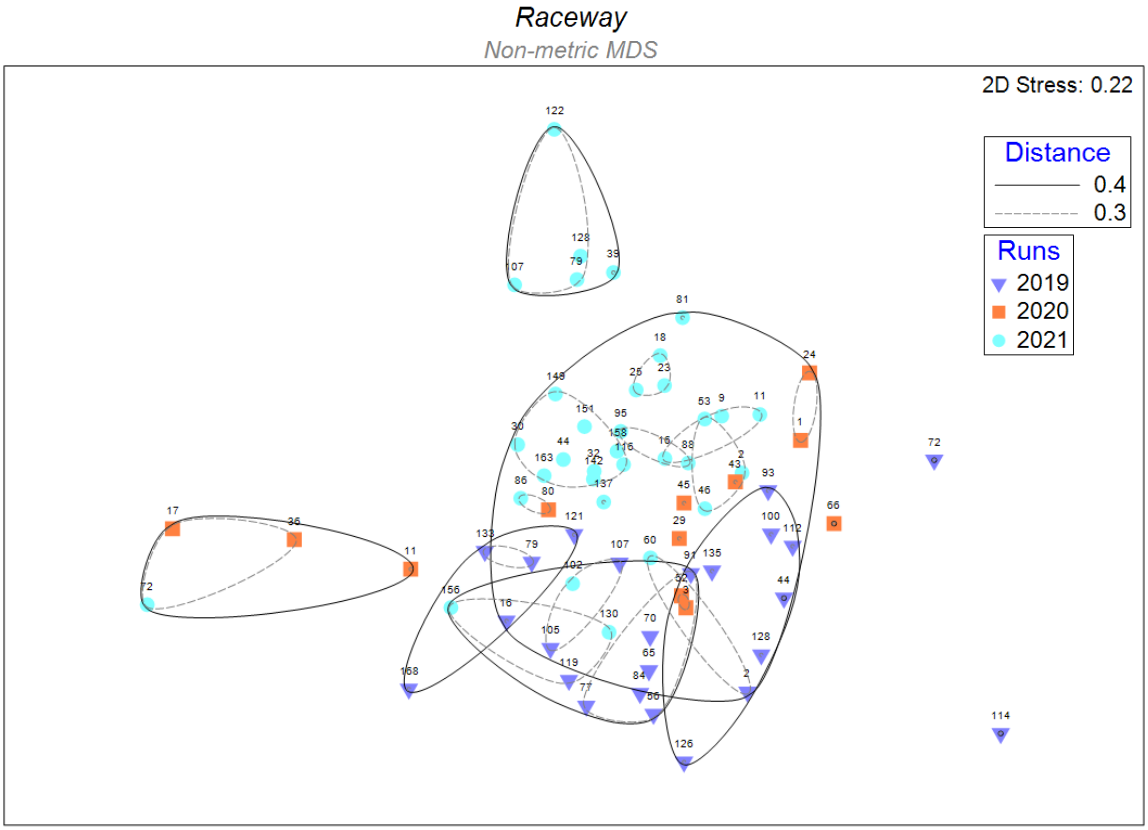

Figure 2.7 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the RW during all three runs*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by run for visualization. Numbers above each symbol indicate the day the sample was taken within its run. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).

A BEST procedure was performed on water quality parameters (temperature, salinity, pH, DO, nitrate, nitrite, and ammonium) in the RW in each run to determine what environmental variables correlate with microbial community composition (Table 9). In 2019, DO had the strongest individual correlation (0.264) with community composition, and nitrate, DO, salinity, and temperature combined had the highest overall correlation (0.320). In 2020, temperature had the highest individual correlation (0.140) and nitrate, nitrite, and temperature combined had the highest overall correlation (0.194). In 2021, no single or combined variables provided a spearman rank correlation coefficient that exceeded 0.006. It is likely the lack of correlation of any variables in 2021 is attributed to the decreased variability of water quality parameters observed in the RW in 2021. With less variation in environmental variables, there is less correlation with community composition.

Table 2.9 *BEST procedure matching the RW's weighted UniFrac distance matrix in correlation with measured water quality parameters per run*

Water quality parameters include temperature, salinity, dissolved oxygen (DO), nitrate (NO3), nitrite (NO2), and ammonia (NH3). Results for runs 2019, 2020, and 2021 span left to right.

BEST Results: Water Quality in the Raceway								
Correlation method: Spearman rank								
Maximum number of variables: 7								
Analyze between: Samples from RW Weighted UniFrac Distance Matrix								
Resemblance measure: D1 Euclidean distance								
Run 2019			*Run 2020*			*Run 2021*		
Number of variables: 1			*Number of variables: 1*			*Number of variables: 1*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.264	DO	1	0.140	Temp	1	0.006	NH3avg
1	0.242	Salinity	1	-0.030	pH	1	-0.017	Temp
1	0.187	Temp	1	-0.074	DO	1	-0.028	Salinity
1	0.039	NH3avg	1	-0.077	NO3avg	1	-0.034	NO3avg
1	-0.002	NO2avg	1	-0.124	Salinity	1	-0.063	DO
1	-0.005	NO3avg	1	-0.131	NO2avg	1	-0.083	pH
1	-0.05	pH	1	-0.223	NH3avg			
Best result for each number of variables			*Best result for each number of variables*			*Best result for each number of variables*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.264	DO	1	0.140	Temp	1	0.006	NH3avg
2	0.295	Salinity, Temp	2	0.190	NO3avg, Temp	2	0.006	NO2avg, NH3avg
3	0.313	NO3avg, Salinity, Temp	3	0.194	NO2avg, NO3avg, Temp	3	-0.006	NO2avg, NO3avg, NH3avg
4	0.32	NO3avg, DO, Salinity, Temp	4	0.188	NO2avg, NO3avg, NH3avg, Temp	4	-0.011	NO2avg, NH3avg, Salinity, Temp
5	0.319	NO3avg, DO, Salinity, pH, Temp	5	0.178	NO2avg, NO3avg, NH3avg, DO, Temp	5	-0.012	NO2avg, NH3avg, Salinity, pH, Temp
6	0.316	NO2avg, NO3avg, DO, Salinity, pH, Temp	6	0.171	NO2avg, NO3avg, NH3avg, DO, pH, Temp	6	-0.018	NO2avg, NO3avg, NH3avg, Salinity, pH, Temp
7	0.256	NO2avg, NO3avg, NH3avg, DO, Salinity,	7	0.022	NO2avg, NO3avg, NH3avg, DO, Salinity,	7	-0.039	NO2avg, NO3avg, NH3avg, DO, Salinity,

Table 2.9 (continued).

Best results			*Best results*			*Best results*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
4	0.320	NO3avg, DO, Salinity, Temp	3	0.194	NO2avg, NO3avg, Temp	1	0.006	NH3avg
5	0.319	NO3avg, DO, Salinity, pH, Temp	3	0.193	NO3avg, NH3avg, Temp	2	0.006	NO2avg, NH3avg
5	0.318	NO2avg, NO3avg, DO, Salinity, Temp	2	0.190	NO3avg, Temp	2	-0.006	NO3avg, NH3avg
6	0.316	NO2avg, NO3avg, DO, Salinity, pH, Temp	4	0.188	NO2avg, NO3avg, NH3avg, Temp	3	-0.006	NO2avg, NO3avg, NH3avg
4	0.314	NO3avg, Salinity, pH, Temp	4	0.187	NO3avg, DO, pH, Temp	2	-0.010	NH3avg, Salinity
3	0.313	NO3avg, Salinity, Temp	4	0.186	NO3avg, NH3avg, DO, Temp	3	-0.010	NO2avg, NH3avg, Salinity
5	0.311	NO2avg, NO3avg, Salinity, pH, Temp	3	0.184	NO3avg, pH, Temp	3	-0.011	NH3avg, Salinity, Temp
4	0.307	NO2avg, NO3avg, Salinity, Temp	4	0.184	NO2avg, NO3avg, pH, Temp	4	-0.011	NO2avg, NH3avg, Salinity, Temp
4	0.305	DO, Salinity, pH, Temp	2	0.183	NH3avg, Temp	4	-0.012	NH3avg, Salinity, pH, Temp
3	0.305	DO, Salinity, Temp	5	0.178	NO2avg, NO3avg, NH3avg, DO, Temp	5	-0.012	NO2avg, NH3avg, Salinity, pH, Temp

Water quality parameters include temperature, salinity, dissolved oxygen (DO), nitrate (NO3), nitrite (NO2), and ammonia (NH3). Results for runs 2019, 2020, and 2021 span left to right.

The PERMANOVA main test on the RW data included in Figure 2.7 is presented in Table 2.10. The results of the analysis indicate neither the run nor the day the sample was collected were significant drivers of community composition. Prior analysis

indicated time in terms of individual runs did shape the RW microbiome (Table 2.8). Clustering of RW samples by run was also evident in the nMDS plots (Figure 2.7). Contradicting results between two PERMANOVAS and visual patterns of orientation observed in the nMDS plot may indicate a complex relationship between time and the RW microbiome that requires further exploration.

Table 2.10 *Permutational multivariate analysis of variance (PERMANOVA) main test performed on the RW to determine differences in the microbiome by runs and day of runs*

The interaction term Runs x Day of Runs was removed by the PERMANOVA due to a lack of sufficient level distribution of the factor "Days of Runs" within levels of factor "Runs" and vice versa. Bolded text indicates a significant result ($p<0.05$).

PERMANOVA Main Test – Raceway						
Sums of squares type: Type III (partial)						
Fixed effects sum to zero for mixed terms						
Permutation method: Permutation of residuals under a reduced model						
Number of permutations: 9999						
Factors						
Name	Type	Levels				
Runs	Fixed	3				
Day of Runs	Random	54				
Excluded terms						
Runs x Day of Runs						
Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Runs	2	0.34859	0.1743	1.939	0.1797	6230
Day of Runs	60	4.9984	0.083306	0.92677	0.7093	9828
Res	6	0.53933	0.089888			
Total	68	6.2328				

A k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels. The algorithm determined which RW samples belonged to one of five clusters, and the results were visualized in an nMDS plot in Figure 2.8. Cluster A contains mostly 2019 samples, B is comprised entirely of 2021 samples in an out group, C includes an out group of 2020 and 2021 samples as well as

two 2019 samples, D is spread across the right side of the plot from top to bottom and includes samples from all three runs, and group E is comprised of primarily 2021 samples.

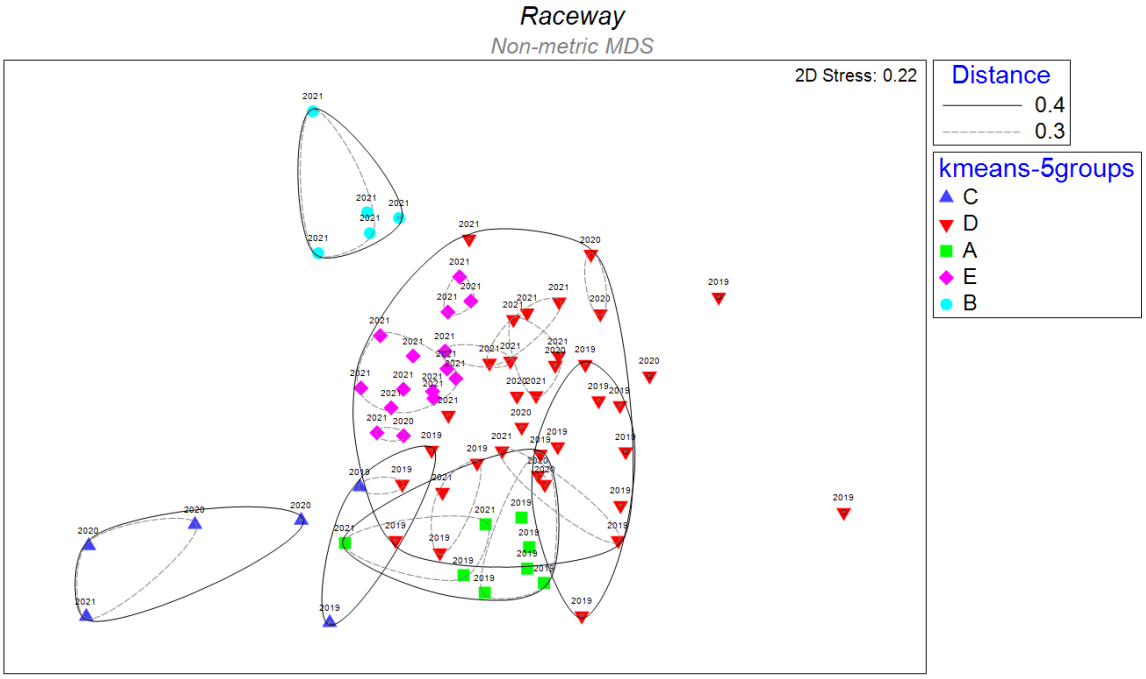

Figure 2.8 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the RW during all three runs organized into k-means clusters*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by five k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).

2.3.6 T Microbiome

The T dataset was visualized via an nMDS plot (Figure 2.9) to determine orientation of samples to each other based on time. This was visualized by coloring samples according to the run they were collected in and labeled by the day of sampling. Samples from 2019 are found in five clusters at a distance of 0.35 (65% similarity) and three clusters at a distance of 0.25 (75% similarity), all of which orient on the right side

of the graph. Three samples from 2019 belong to no group. A cluster at a distance of 0.35 on the left side of the graph includes the majority of 2020 and 2021 samples. Most 2020 samples form a group clustered at a distance of 0.25. Most 2021 samples form three clusters at a distance of 0.25. There are clear distinctions between samples taken in 2019, 2020, and 2021 in the T.

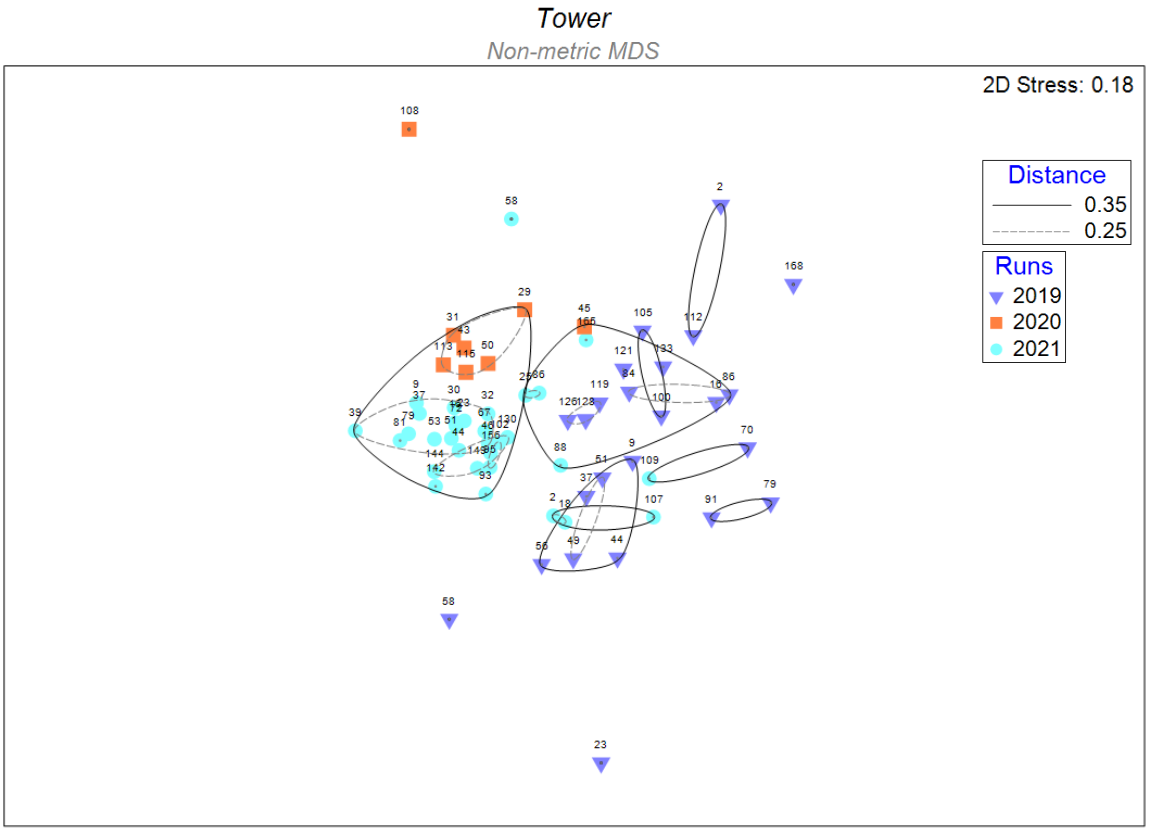

Figure 2.9 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the T during all three runs*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by runs for visualization. Numbers above each symbol indicate the day the sample was taken within its run. A CLUSTER analysis generated distance contour lines at 0.35 (65% similar) and 0.25 (75% similar).

A BEST procedure was performed using water quality parameters (temperature, salinity, pH, DO, nitrate, nitrite, and ammonium) in the T in each run to determine what

environmental variables correlate with microbial community composition (Table 2.11). In 2019, ammonium had the strongest individual correlation (0.030) with community composition, but no other individual variable provided a spearman rank correlation coefficient that exceeded 0.000. Ammonium and salinity combined had the highest overall correlation (0.015). In 2020, nitrate had the highest individual correlation (0.360), and nitrate, nitrite, and salinity combined had the highest overall correlation (0.434). In 2021, nitrate had the highest individual correlation (0.161), and nitrate and nitrite had the highest combined correlation (0.161) in 2021. Lower spearman rank correlation coefficients in 2019 and 2021 compared to 2020 reflect the higher variability in parameters observed in 2020 in the T compared to the other two runs.

Table 2.11 *BEST procedure matching the T's weighted UniFrac distance matrix in correlation with measured water quality parameters per run*

Water quality parameters include temperature, salinity, dissolved oxygen (DO), nitrate (NO3), nitrite (NO2), and ammonia (NH3). Results for runs 2019, 2020, and 2021 span left to right.

BEST Results: Water Quality in the Tower								
Correlation method: Spearman rank								
Maximum number of variables: 7								
Analyze between: Samples from T Weighted UniFrac Distance Matrix								
Resemblance measure: D1 Euclidean distance								
Run 2019			*Run 2020*			*Run 2021*		
Number of variables: 1			*Number of variables: 1*			*Number of variables: 1*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.030	NH3avg	1	0.360	NO3avg	1	0.161	NO3avg
1	-0.011	Salinity	1	0.315	Salinity	1	0.052	Salinity
1	-0.071	pH	1	0.304	Temp	1	0.033	Temp
1	-0.083	DO	1	0.184	pH	1	-0.126	pH
1	-0.126	Temp	1	-0.064	NO2avg	1	-0.154	DO
1	-0.158	NO2avg	1	-0.155	NH3avg	1	-0.211	NH3avg
1	-0.186	NO3avg	1	-0.156	DO			
Best result for each number of variables			*Best result for each number of variables*			*Best result for each number of variables*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.030	NH3avg	1	0.360	NO3avg	1	0.161	NO3avg
2	0.015	NH3avg, Salinity	2	0.434	NO3avg, Salinity	2	0.161	NO2avg, NO3avg
3	0.015	NH3avg, Salinity, pH	3	0.434	NO2avg, NO3avg, Salinity	3	0.093	NO2avg, NO3avg, pH
4	-0.007	NO2avg, NH3avg, Salinity, pH	4	0.434	NO2avg, NO3avg, Salinity, pH	4	0.082	NO2avg, NO3avg, pH, Temp
5	-0.041	NO2avg, NH3avg, DO, Salinity, pH	5	0.418	NO2avg, NO3avg, NH3avg, Salinity, pH	5	0.073	NO2avg, NO3avg, NH3avg, pH, Temp
6	-0.076	NO2avg, NH3avg, DO, Salinity, pH, Temp	6	0.337	NO2avg, NO3avg, NH3avg, DO, Salinity, Temp	6	0.057	NO2avg, NO3avg, NH3avg, Salinity, pH, Temp
7	-0.153	NO2avg, NO3avg, NH3avg, DO, Salinity, pH, Temp	7	0.337	NO2avg, NO3avg, NH3avg, DO, Salinity, pH, Temp	7	-0.056	NO2avg, NO3avg, NH3avg, DO, Salinity, pH, Temp

Table 2.11 (continued).

Best results			*Best results*			*Best results*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.030	NH3avg	2	0.434	NO3avg, Salinity	1	0.161	NO3avg
2	0.015	NH3avg, Salinity	3	0.434	NO2avg, NO3avg, Salinity	2	0.161	NO2avg, NO3avg
3	0.015	NH3avg, Salinity, pH	3	0.434	NO3avg, Salinity, pH	2	0.093	NO3avg, pH
3	-0.002	NO2avg, NH3avg, Salinity	4	0.434	NO2avg, NO3avg, Salinity, pH	3	0.093	NO2avg, NO3avg, pH
4	-0.007	NO2avg, NH3avg, Salinity, pH	3	0.418	NO3avg, NH3avg, Salinity	2	0.086	NO3avg, NH3avg
2	-0.009	Salinity, pH	4	0.418	NO2avg, NO3avg, NH3avg, Salinity	3	0.086	NO2avg, NO3avg, NH3avg
1	-0.011	Salinity	4	0.418	NO3avg, NH3avg, Salinity, pH	3	0.082	NO3avg, pH, Temp
2	-0.017	NO2avg, Salinity	5	0.418	NO2avg, NO3avg, NH3avg, Salinity, pH	4	0.082	NO2avg, NO3avg, pH, Temp
4	-0.021	NH3avg, DO, Salinity, pH	3	0.365	NO2avg, NO3avg, pH	2	0.073	NO3avg, Temp
3	-0.021	NH3avg, DO, Salinity	2	0.361	NO3avg, pH	3	0.073	NO2avg, NO3avg, Temp

Water quality parameters include temperature, salinity, dissolved oxygen (DO), nitrate (NO3), nitrite (NO2), and ammonia (NH3). Results for runs 2019, 2020, and 2021 span left to right.

The PERMANOVA main test on the T data included in Figure 2.9 is presented in Table 2.12. The results of the analysis indicate the run the sample was collected during was a significant driver of community composition, but the day of sampling was not. This agrees with prior analysis indicating all three runs in the T had distinct communities

(Table 2.8). This distinction between runs was also visually evident in the nMDS plot, where samples clustered by run (Figure 2.9).

Table 2.12 *Permutational multivariate analysis of variance (PERMANOVA) main test performed on the T to determine differences in the microbiome by runs and day of runs*

The interaction term Runs x Day of Runs was removed by the PERMANOVA due to a lack of sufficient level distribution of the factor "Days of Runs" within levels of factor "Runs" and vice versa. Bolded text indicates a significant result ($p<0.05$).

PERMANOVA Main Test - Tower						
Sums of squares type: Type III (partial)						
Fixed effects sum to zero for mixed terms						
Permutation method: Permutation of residuals under a reduced model						
Number of permutations: 9999						
Factors						
Name	Type	Levels				
Runs	Fixed	3				
Day of Runs	Random	54				
Excluded terms						
Runs x Day of Runs						
Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Runs	2	0.36338	0.36338	5.2561	**0.0007**	9673
Day of Runs	52	3.0054	0.057796	0.83599	0.902	9801
Res	9	0.62222	0.069135			
Total	63	4.6366				

As was done for the RW, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels in the T. The algorithm organized samples into five clusters, and the results were visualized in an nMDS plot in Figure 2.10. Cluster A contains mostly 2019 and 2021 samples, B is comprised entirely of 2021 samples, C includes mostly 2019 and 2021 samples, D is comprised entirely of 2020 samples, and group E is comprised entirely of 2021 samples.

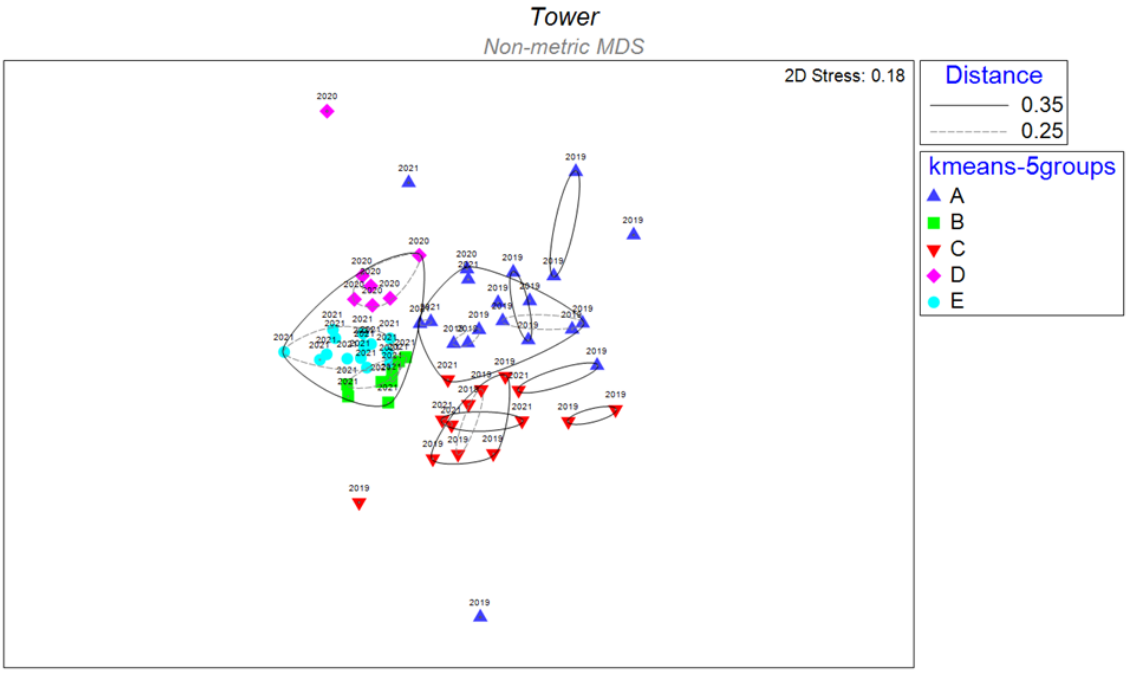

Figure 2.10 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the T during all three runs organized into k-means clusters*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by five k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.35 (65% similar) and 0.25 (75% similar).

2.3.7 LT Microbiome

The LT dataset was visualized via an nMDS plot (Figure 2.11) to determine orientation of samples to each other based on time. This was visualized by coloring samples according to the run they were collected in and labeled by the day of sampling. There are no discernable patterns observed between samples and the runs they were collected during. There is also no pattern by day of sampling. Some samples from 2021 form clusters at a distance of 0.2 (80% similarity), but most 2021 samples spread across the graph amongst samples from 2019 and 2020.

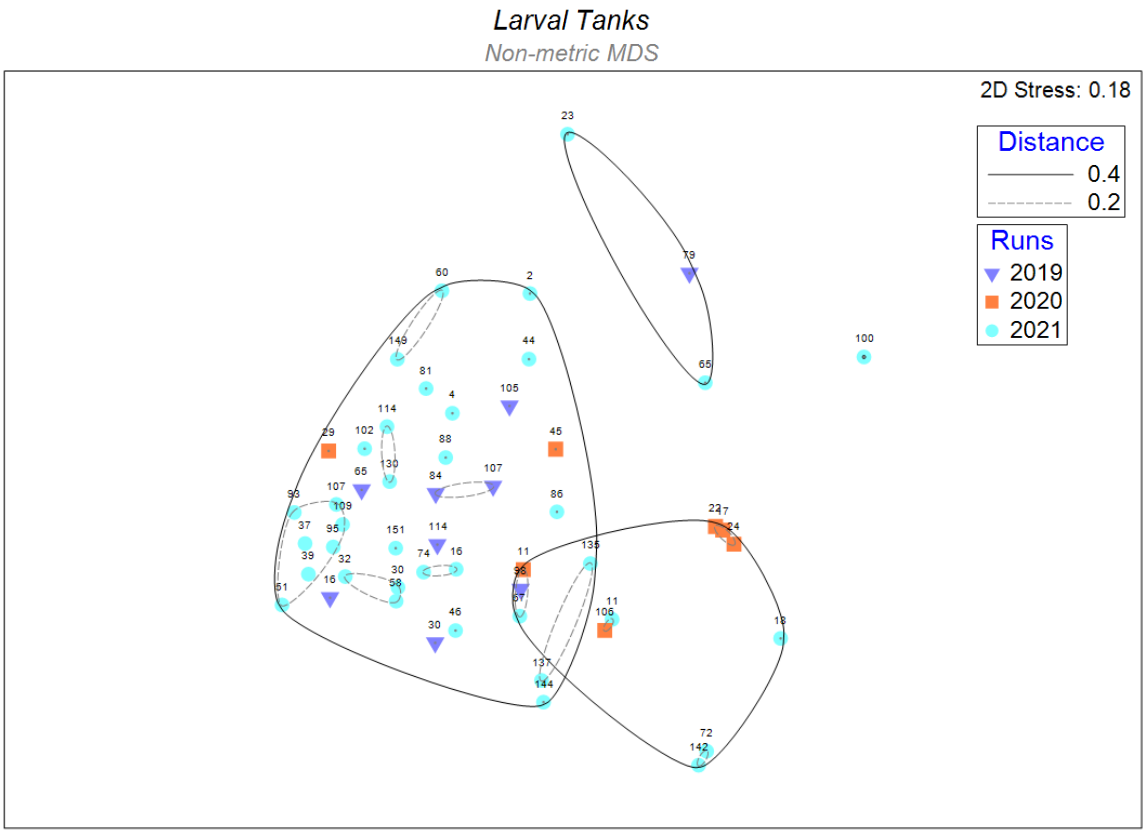

Figure 2.11 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the LT during all three runs*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by runs for visualization. Numbers above each symbol indicate the day the sample was taken within its run. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

The PERMANOVA main test on the LT data included in Figure 2.11 is presented in Table 2.13. The results of this analysis indicate neither the run the sample was collected during nor the day of sampling were significant drivers of community composition. Prior analysis indicated time in terms of individual runs did shape the LT microbiome (Table 2.8), but a lack of clustering of LT samples in the nMDS plot agrees time is not a driving factor of community composition in the LT microbiome (Figure 2.11).

Table 2.13 *Permutational multivariate analysis of variance (PERMANOVA) main test performed on the LT to determine differences in the microbiome by runs and day of runs*

The interaction term Runs x Day of Runs was removed by the PERMANOVA due to a lack of sufficient level distribution of the factor "Days of Runs" within levels of factor "Runs" and vice versa. Bolded text indicates a significant result ($p<0.05$).

PERMANOVA Main Test – Larval Tank						
Sums of squares type: Type III (partial)						
Fixed effects sum to zero for mixed terms						
Permutation method: Permutation of residuals under a reduced model						
Number of permutations: 9999						
Factors						
Name	Type	Levels				
Runs	Fixed	3				
Day of Runs	Random	54				
Excluded terms						
Runs x Day of Runs						
Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Runs	2	0.099116	0.049558	0.80232	0.4545	3364
Day of Runs	45	3.5851	0.079669	1.2898	0.2445	9906
Res	4	0.24707	0.061768			
Total	51	4.1874				

As was done for the RW and T, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels in the LT. The algorithm organized samples into five clusters, and the results were visualized in an nMDS plot in Figure 2.12. Cluster A contains mostly 2021 samples, B contains 2020 and 2021 samples, C includes mostly 2021 samples, D is a small cluster of entirely 2020 samples, and group E is comprised mostly of 2021 samples.

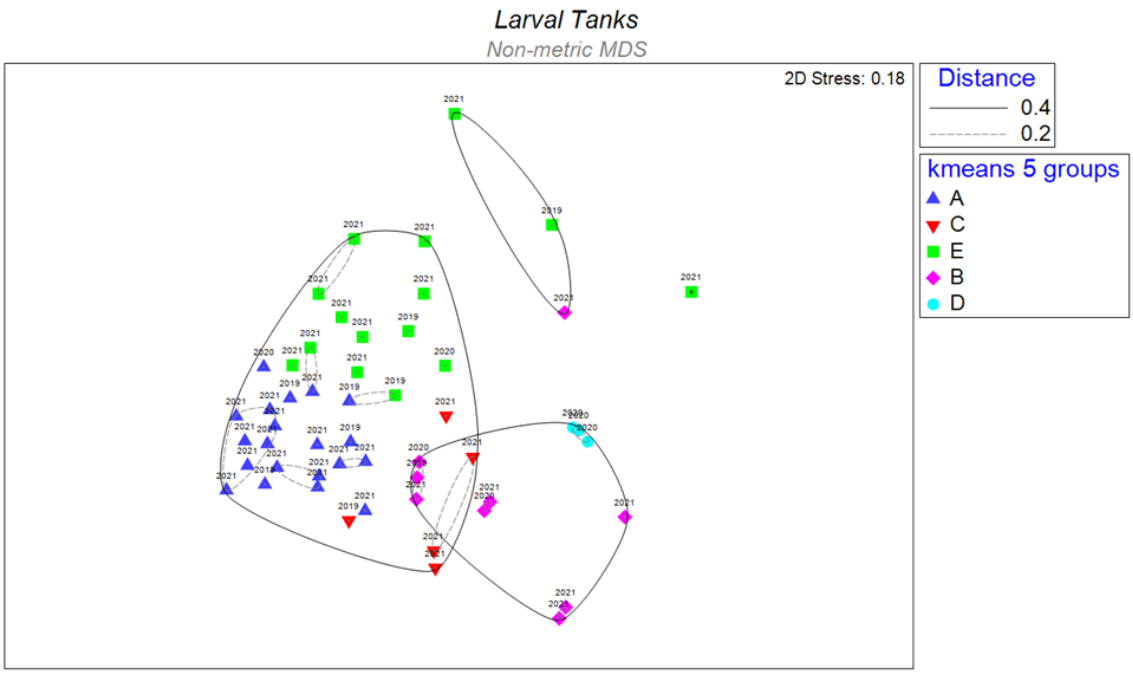

Figure 2.12 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the LT during all three runs organized into k-means clusters*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by five k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

2.3.8 AH Microbiome

The AH dataset was visualized via an nMDS plot (Figure 2.13) to determine orientation of samples to each other based on time. This was visualized by coloring samples according to the run they were collected in and labeled by the day of sampling. There are no discernable patterns observed between samples and the runs they were collected during. There is also no clear pattern by day of sampling. Samples from 2019, 2020, and 2021 are interspersed throughout the plot. Most clusters drawn at distances of 0.4 (60% similarity) are comprised of samples from all three runs. Some samples taken one to two days apart from one another group together, which is likely due to the AH

being sampled on subsequent days (AH1 on one day, AH2 24 hours later). The lack of clustering is likely due to the AH contents being replaced every other day.

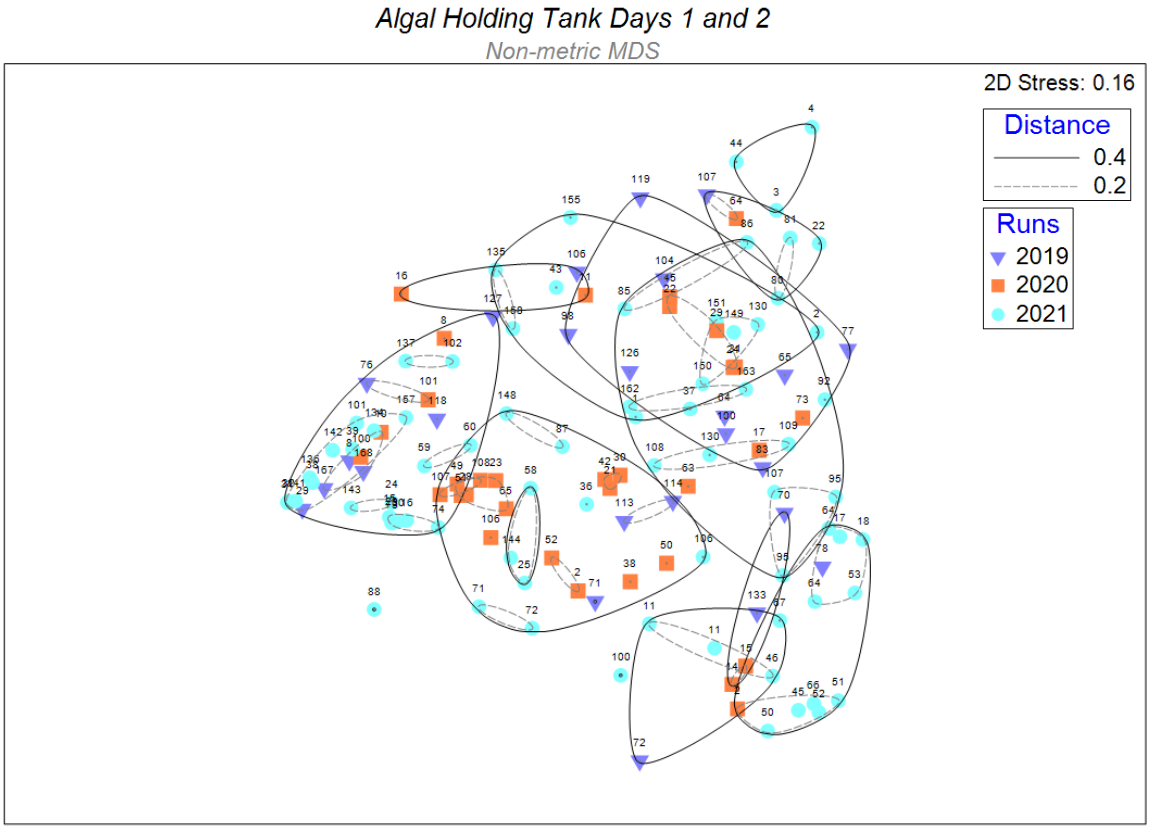

Figure 2.13 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the AH during all three runs*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by runs for visualization. Numbers above each symbol indicate the day the sample was taken within its run. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

The PERMANOVA main test on the AH data included in Figure 2.13 is presented in Table 2.14. The results of the analysis indicate the run the sample was collected during and the day of sampling were significant drivers of community composition. This agrees with prior analysis (Table 2.8) indicating runs shaped the AH microbiome. However, clustering of AH samples by run was not evident in the nMDS plots (Figure 2.13). The

interaction term having the higher Pseudo-F value (F = 8.5242) suggests there are differences between the days of sampling within each run that rival the main effects of each individual term. This is corroborated by the fact the AH is replaced every other day and exists outside the RAS loop.

Table 2.14 *Permutational multivariate analysis of variance (PERMANOVA) main test performed on the AH to determine differences in the microbiome by runs and day of runs*

Bolded text indicates a significant result ($p<0.05$).

PERMANOVA Main Test - Algal Holding Tank						
Sums of squares type: Type III (partial)						
Fixed effects sum to zero for mixed terms						
Permutation method: Permutation of residuals under a reduced model						
Number of permutations: 9999						
Factors						
Name	Type	Levels				
Runs	Fixed	3				
Day of Runs	Random	54				
Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Runs	2	0.83698	0.41849	2.8391	**0.0072**	9932
Day of Runs	96	13.583	0.14149	8.0772	**0.0001**	9908
Runs x Day of Runs	29	4.3304	0.14933	8.5242	**0.0037**	9944
Res	5	0.087589	0.017518			
Total	132	19.44				

As was done for the within-loop compartments, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels in the AH. The algorithm organized samples into five clusters, and the results were visualized in an nMDS plot in Figure 2.14. All five clusters (A through E) are comprised of samples from all three runs. Clusters A, B, D, and E are all loosely grouped and spread amongst each other. Cluster C is more tightly knit. Though clusters were identified by the k-means algorithm, time does not appear to be a factor driving community composition.

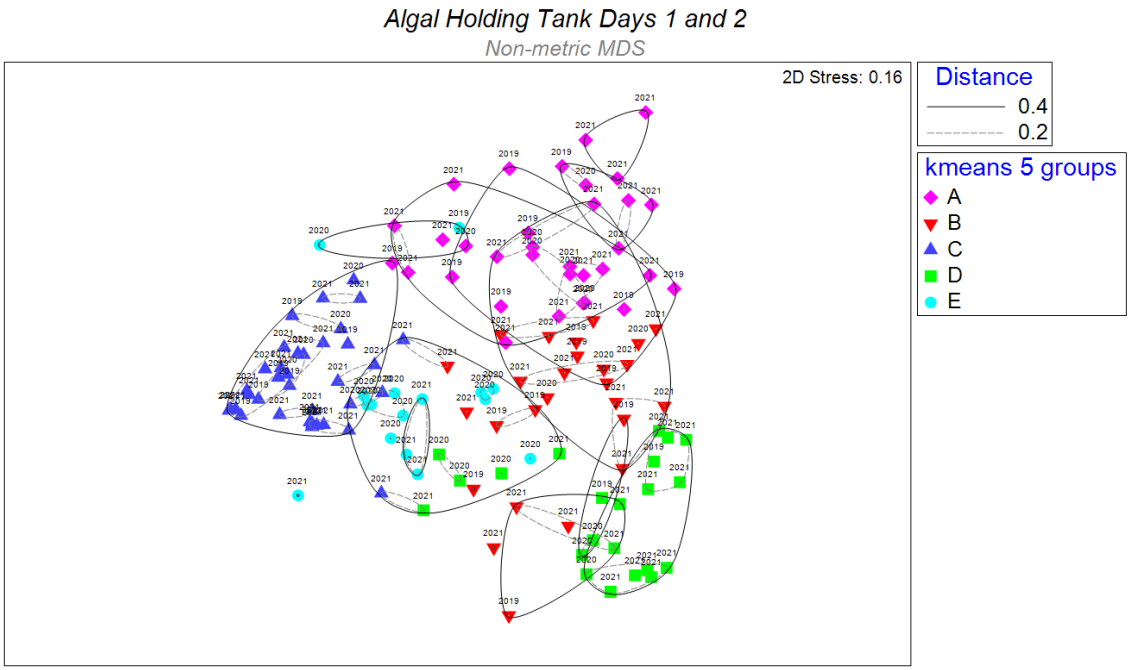

Figure 2.14 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the AH during all three runs organized into k-means clusters*

. Data derived from the weighted UniFrac distance matrix. Samples are grouped by five k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

2.3.9 RAS Microbiome Stability

The moving window analysis (MWA) performed on each compartment offers a means of quantifying stability of a microbiome in terms of UniFrac dissimilarity. Distances closer to 0 indicates paired samples are highly alike (0 distance equates to 100% similarity), so high stability is quantified as low distances over time and vice versa.

Results of the MWA for the RW are shown as barplots detailing the distances between successive samples (Figure 2.15), distances between the first timepoint in 2019 and all other samples (Figure 2.16), and successive distances smoothed using the LOESS technique (Figure 2.17).

In the RW, community composition between successive samples was, on average, at a distance of 0.39 ± 0.09 in 2019, 0.43 ± 0.08 in 2020, and 0.34 ± 0.10 in 2021 (Figure 2.15). In terms of percent similarity in microbial composition, successive samples in 2019 were on average 61.41 ± 9.21% similar, samples in 2020 were 56.66 ± 8.35% similar, and samples in 2021 were 65.55 ± 10.23% similar. Using a two-sample, two-tailed t-test assuming unequal variance, average stability in 2019 was not significantly different than the average stabilities of 2020 and 2021 in the RW ($p=0.15$, $p=0.12$, respectively). The difference between average stability in 2020 and 2021 was significantly different ($p=0.01$). These results are consistent with previous PERMANOVA results in the RW (Table 8) and suggest the changes in community composition between 2020 and 2021 coincided with an increase in average stability between the two runs.

For distances between successive samples (Figure 2.15), the lowest distance between successive samples (highest moment of stability) occurred in 2019 between days 56 and 65 (distance of 0.17, 82.82% similar). The highest distance between successive samples (highest moment of instability) occurred in 2021 between days 72 and 79 (distance of 0.64, 36.03% similar). In 2019, the RW microbiome was initially unstable until day 56. Stability lowered after day 70, increased over time, then lowered again after day 114. In 2020, microbial stability was initially near average for 2020 until a spike occurred between days 17 and 24. Stability generally increased between days 24 and 45, but the run ended at above average instability from day 45 onward. In 2021, microbial stability was initially above average. It generally remained stable until a spike in

instability occurred between days 60 and 79. Following this spike, a pattern appears to have emerged of momentary instability followed by increased stability.

For distances between the first timepoint in 2019 (Figure 2.16), the starting time point in the RW was day 2 in 2019. On average, stability in terms of the experimental start point was a distance of 0.44 ± 0.04 in 2019, 0.46 ± 0.07 in 2020, and 0.45 ± 0.06 in 2021 for an overall average of 0.45 ± 0.06. Using a two-sample, two-tailed t-test assuming unequal variance, average stabilities from the starting time in all three runs were not significantly different (2019-2020: $p=0.47$; 2020-2021: $p=0.51$; 2019-2021: $p=0.93$). This result suggests there was no significant difference in RW microbial stability across all three runs.

The lowest distance from the run start occurred in 2021 on day 60 (distance of 0.28) (Figure 2.16). The highest distance occurred in 2020 on day 17 (distance of 0.61). In 2019, distances remained at about 0.40 to 0.45 throughout the run with the exception of brief rises to 0.51 on days 72, 114, and 168. This mostly coincides with trends in stability over time (Figure 2.15) in 2019 where instability occurred between days 70 and 72 as well as days 112 and 119. In 2020, distances rose from 0.48 on day 1 to 0.61 on day 17 (the highest moment of instability). Distances then lowered to 0.37 by day 29, rose to 0.53 on day 36, then lowered to a 0.35-0.45 for the remainder of the run. The high instability on day 17 coincides with a spike in stability over time (Figure 2.15) between days 17 and 24 in 2020. In 2021, distances were generally 0.40-0.45 until a rise to 0.52 on day 39 followed by a fall to 0.28 by day 60 (the highest moment of stability). Distances then rose to 0.60 the subsequent time point, lowered to 0.32 by day 86, rose again to 0.55 by day 122, then remained at 0.38-0.49 for the rest of the run. The rise in

instability near day 72 coincides with the highest moment of instability in Figure 2.15, as does the brief rise in stability on day 60.

The LOESS smoothing technique permits clearer analysis of regime shifts in timescale data. In Figure 2.17, the RW data was smoothed at a span of 20% to maximize the number of regime shifts detected. Figure 2.17 shows the distances between successive samples with the smoothing technique applied and regime shifts detected colored.

Smoothing stability data over time, displayed in Figure 2.17, showcased general trends in rising and falling stability already identified in Figure 2.15. In 2019, stability lowered between days 70 and 77 and then again between days 106 and 119. In 2020, stability lowered over time to peak between days 24 and 29 before rising for the remainder of the run. In 2021, stability decreased between days 18 and 79 before mostly increasing throughout the rest of the run. STARS detected regime shifts on days 66 to 80 in 2020, 45 to 53 in 2021, and 129 onward in 2021. The shift at the end of 2020 indicates a change between 2020's end and 2021's start. The shift on days 45 to 53 in 2021 prelude the rise in instability peaking by days 70 to 72. The final detected shifts occur on the last eight time points in the experiment in the RW, suggesting the observed increases in stability at the end of the experiment were significant.

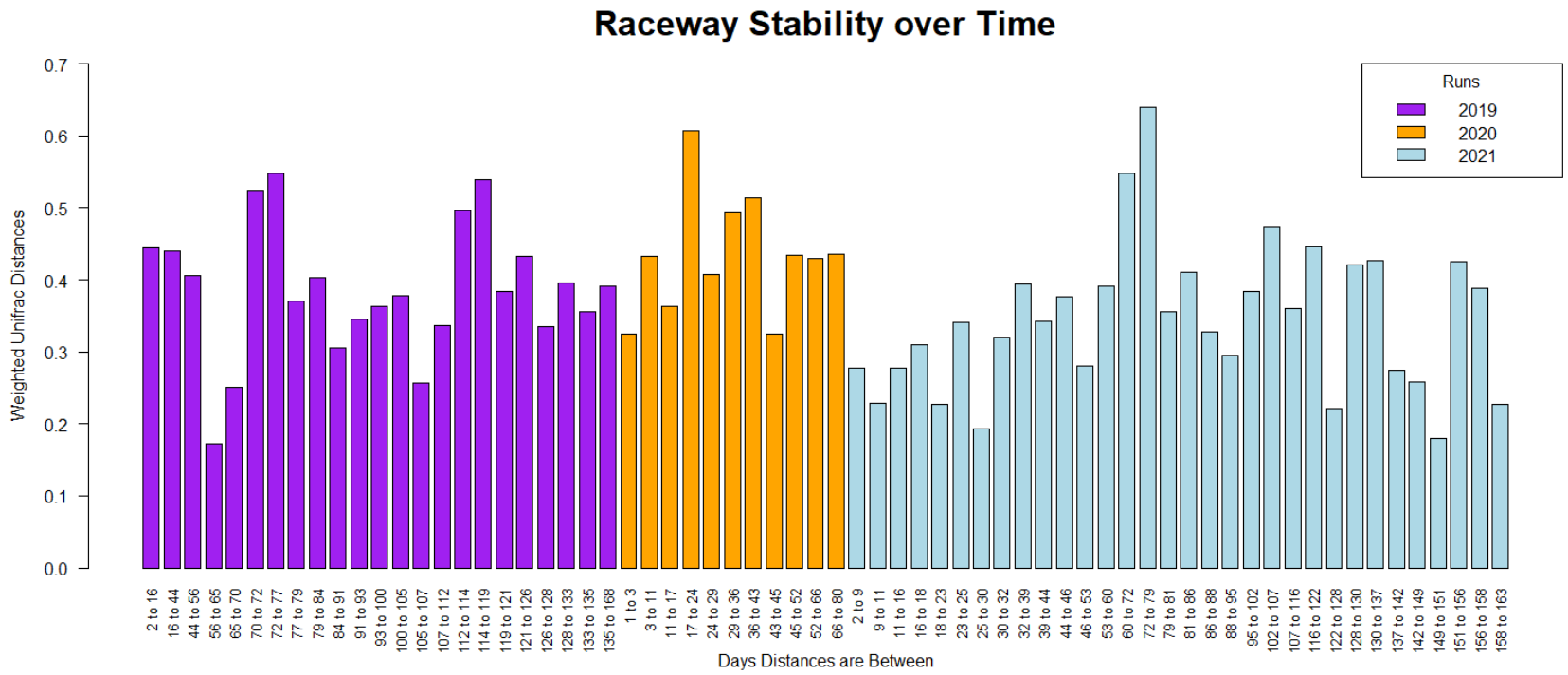

Figure 2.15 *Moving window analysis applied to successive weighted UniFrac distances for the RW*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

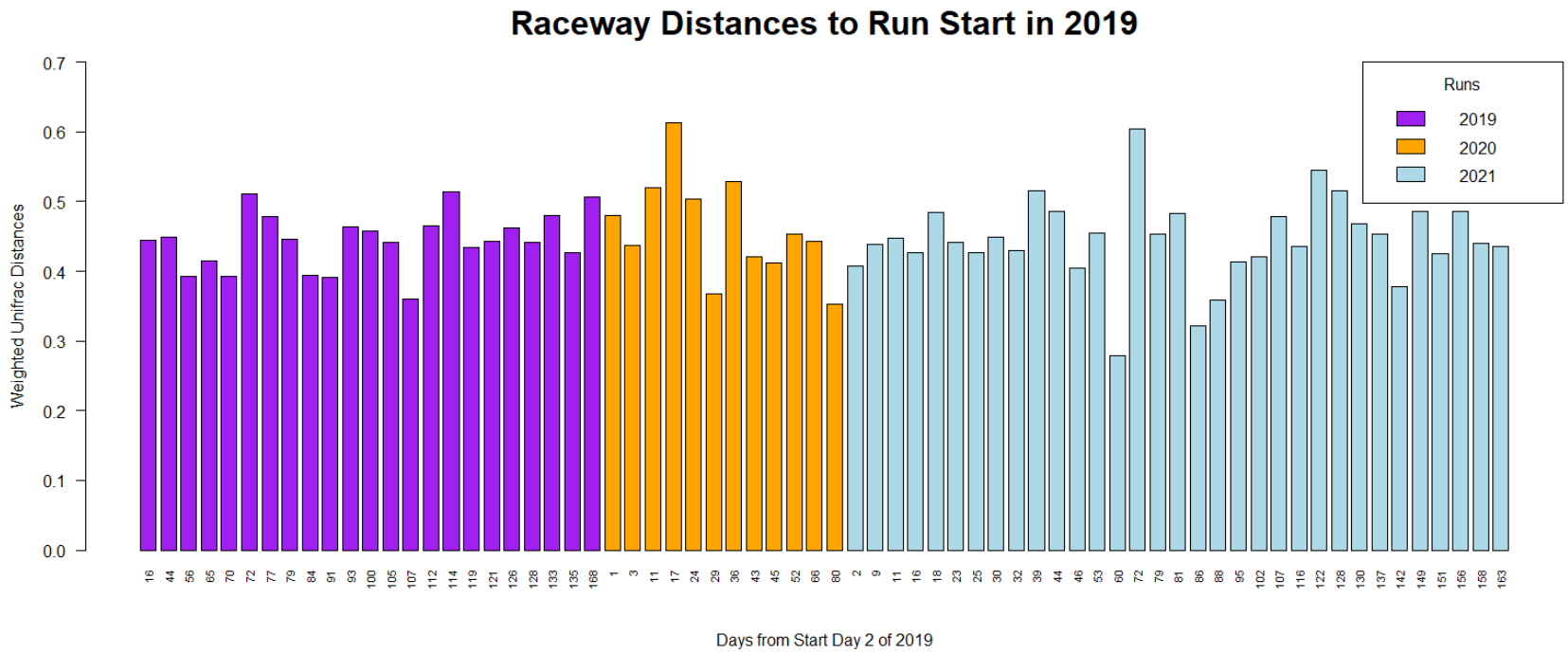

Figure 2.16 *Moving window analysis applied to the weighted UniFrac distances between the first sampling day and all subsequent samples in the RW*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between the beginning of the experiment (day 2 of 2019) and all subsequent samples (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

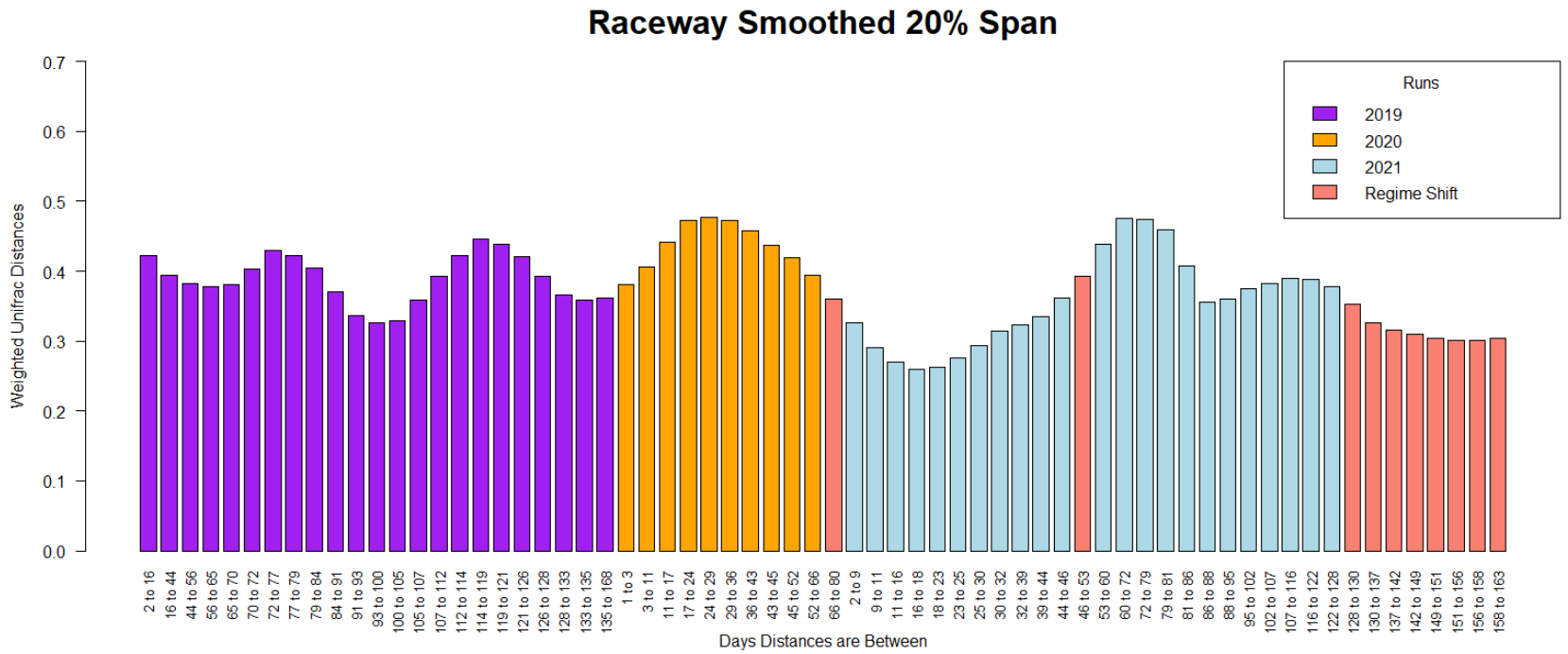

Figure 2.17 *Moving window analysis applied to successive weighted UniFrac distances for the RW smoothed using LOESS*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). Regime shifts detected by the Student's t-test Analysis of Regime Shifts (STARS) are colored salmon. UniFrac distances were "smoothed" using Local Regression (LOESS). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

Results of the MWA for the T are shown as barplots detailing the distances between successive samples (Figure 2.18), distances between the first timepoint in 2019 and all other samples (Figure 2.19), and successive distances smoothed using the LOESS technique (Figure 2.20).

In the T, community composition between successive samples was, on average, at a distance of 0.35 ± 0.08 in 2019, 0.26 ± 0.11 in 2020, and 0.28 ± 0.07 in 2021 (Figure 2.18). In terms of percent similarity in microbial composition, successive samples in 2019 were on average 65.40 ± 7.53% similar, samples in 2020 were 73.79 ± 11.32% similar, and samples in 2021 were 72.13 ± 7.18% similar. Using a two-sample, two-tailed t-test assuming unequal variance, average stability in 2020 was not significantly different than the average stabilities of 2019 and 2021 in the T ($p=0.10$, $p=0.72$, respectively). The difference between average stability in 2019 and 2021 was significantly different ($p<0.01$). These results suggest changes in stability in the T occur over a two year period.

For distances between successive samples (Figure 2.18), the lowest distance between successive samples (highest moment of stability) occurred in 2020 between days 31 and 43 (distance of 0.095, 90.51% similar) The highest distance between successive samples (highest moment of instability) occurred in 2019 between days 16 and 23 (distance of 0.48, 51.91% similar). In 2019, the T microbiome experienced a spike of instability between days 16 and 37 before lowering. A general rise in instability occurred from day 49 to 91 then declined until day 28. 2019 ended in a period of above average instability. In 2020, microbial stability was initially high (below average distances). Between days 43 and 113, stability was lower than average. At the end of the run, there

was a period of high stability. In 2021, microbial stability was initially below average then rose and fell periodically over the course of the run.

For distances between the first timepoint in 2019 (Figure 2.19), the starting time point in the T was day 2 in 2019. On average, stability in terms of the experimental start point was a distance of 0.45 ± 0.06 in 2019, 0.49 ± 0.06 in 2020, and 0.48 ± 0.03 in 2021 for an overall average of 0.47 ± 0.05. Using a two-sample, two-tailed t-test assuming unequal variance, average stabilities from the starting time in 2019 and 2021 were significant different ($p=0.04$) while averages were not significantly different between 2019 and 2020 or 2020 and 2021 ($p=0.12$, $p=0.64$, respectively). This result is consistent with results viewing stability over time (Figure 2.18), indicating that stability in terms of each sample to the experimental start changed slowly over time.

The lowest distance from the run start occurred in 2019 on day 112 (distance of 0.34) (Figure 2.19). The highest distance occurred in 2020 on day 108 (distance of 0.59). In 2019, distances were initially about 0.35 until rising to 0.58 by day 23. Distances generally rose to a peak of 0.43 by day 121 with the exception of a brief drop to 0.34 on day 112 (the highest moment of stability). Distances then lowered to 0.40 by day 128 before generally rising for the rest of the run to 0.51 on day 168. The rise in instability day 23 coincides with peaks in stability over time (Figure 2.18) between days 16 and 37. The highest moment of stability observed on day 112 does not, however, correspond to a similar rise in stability in the same general time when viewing stability over time (Figure 2.18). In 2020, distances were initially at 0.49-0.50 until falling to 0.39 on day 45. Distances then rose to 0.59 by day 108 (the highest moment of instability) before lowering to 0.48 by day 115. The rise in stability on day 45 slightly coincides with a peak

observed in Figure 2.18 between days 31 and 43, though stability between successive samples was lower between days 45 and 119. In 2021, distances generally ranged between 0.43 and 0.50 throughout the run. Momentary peaks of 0.54-0.55 occurred on days 39 and 81. Trends in stability in terms of experimental start (Figure 2.19) conflict with trends observed through successive distances (Figure 2.18). Conflictions between stability quantification methods in the T are especially apparent in 2021, where clear spikes in stability and instability seen in Figure 2.18 are not observed in Figure 2.19.

The LOESS smoothing technique permits clearer analysis of regime shifts in timescale data. In Figure 2.20, the T data was smoothed at a span of 15% to maximize the number of regime shifts detected. Figure 2.20 shows the distances between successive samples with the smoothing technique applied and regime shifts detected colored.

Smoothing stability data over time, displayed in Figure 2.20, showcased general trends in rising and falling stability already identified in Figure 2.18. In 2019, stability lowered between days 9 and 23 and then again between days 51 and 91. Stability rose until days 119 to 121, after which stability lowered for the remainder of the run. In 2020, stability mostly lowered until days 108 to 119 where stability then rose for the remainder of the run. In 2021, five valleys (periods of high stability) occurred between days 23 and 39, 46 and 81, 67 and 81, 88 and 107, and 142 and 165. STARS detected regime shifts on days 119 to 121 in 2019 and days 142 onward in 2021. The detected shift in 2019 matches the shift to above average stability starting between days 112 and 119. The final detected shifts occur on the last four time points in the experiment in the T, suggesting the observed rise in stability at the end of the experiment was significant.

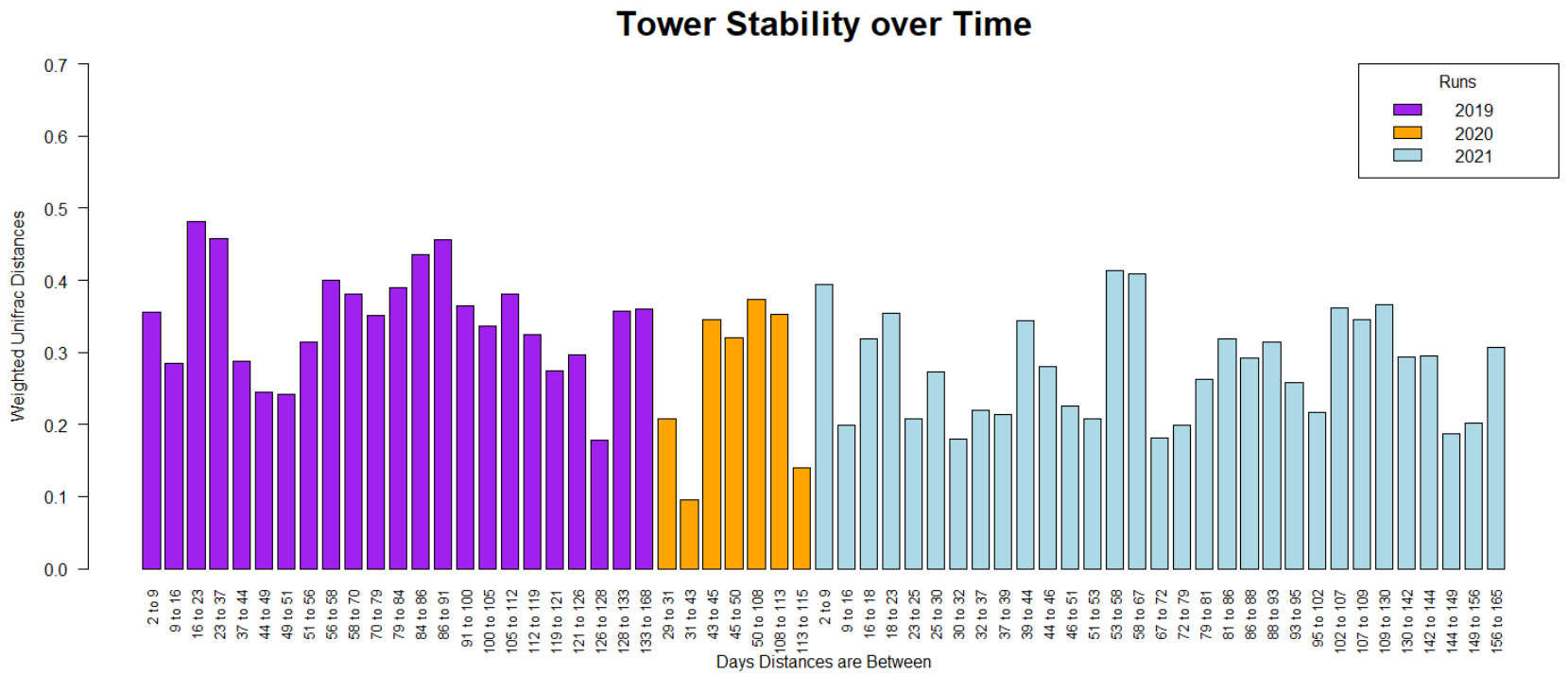

Figure 2.18 *Moving window analysis applied to successive weighted UniFrac distances for the T*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

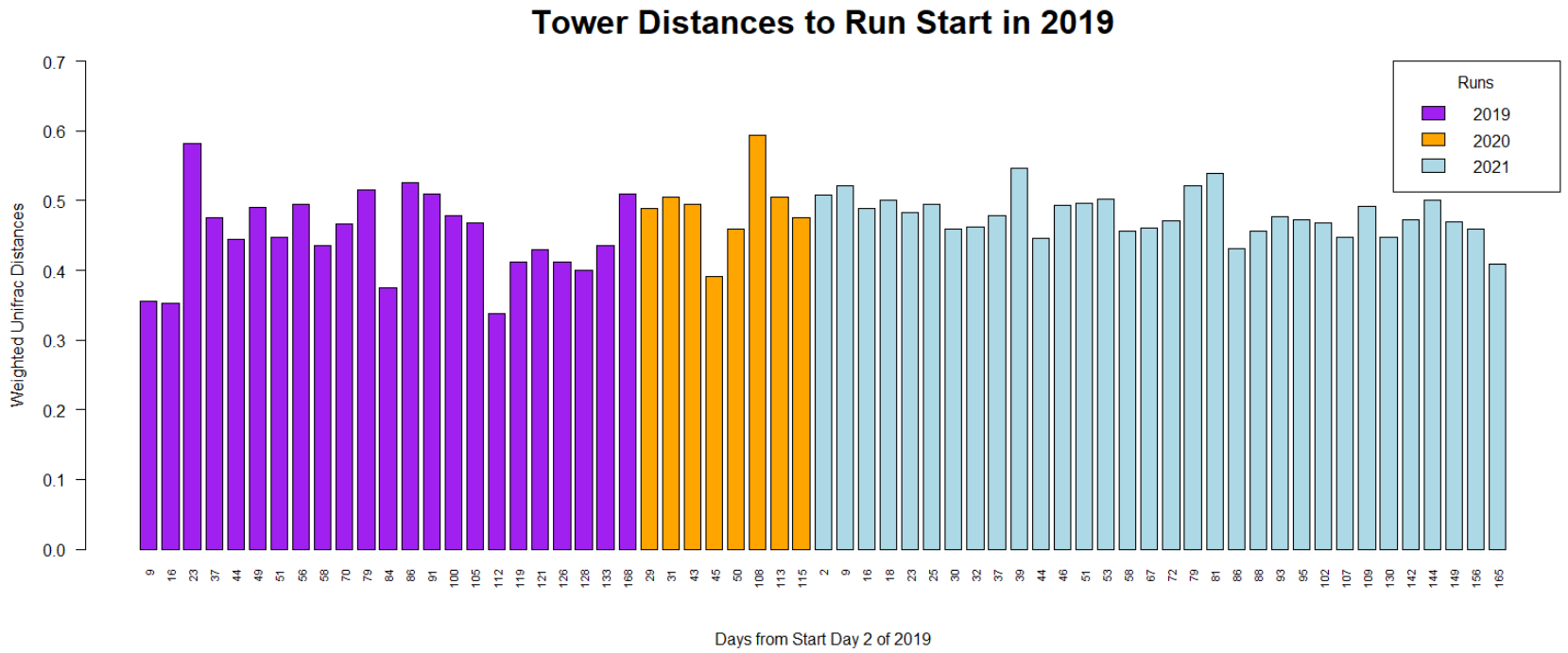

Figure 2.19 *Moving window analysis applied to the weighted UniFrac distances between the first sampling day and all subsequent samples in the T*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between the beginning of the experiment (day 2 of 2019) and all subsequent samples (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

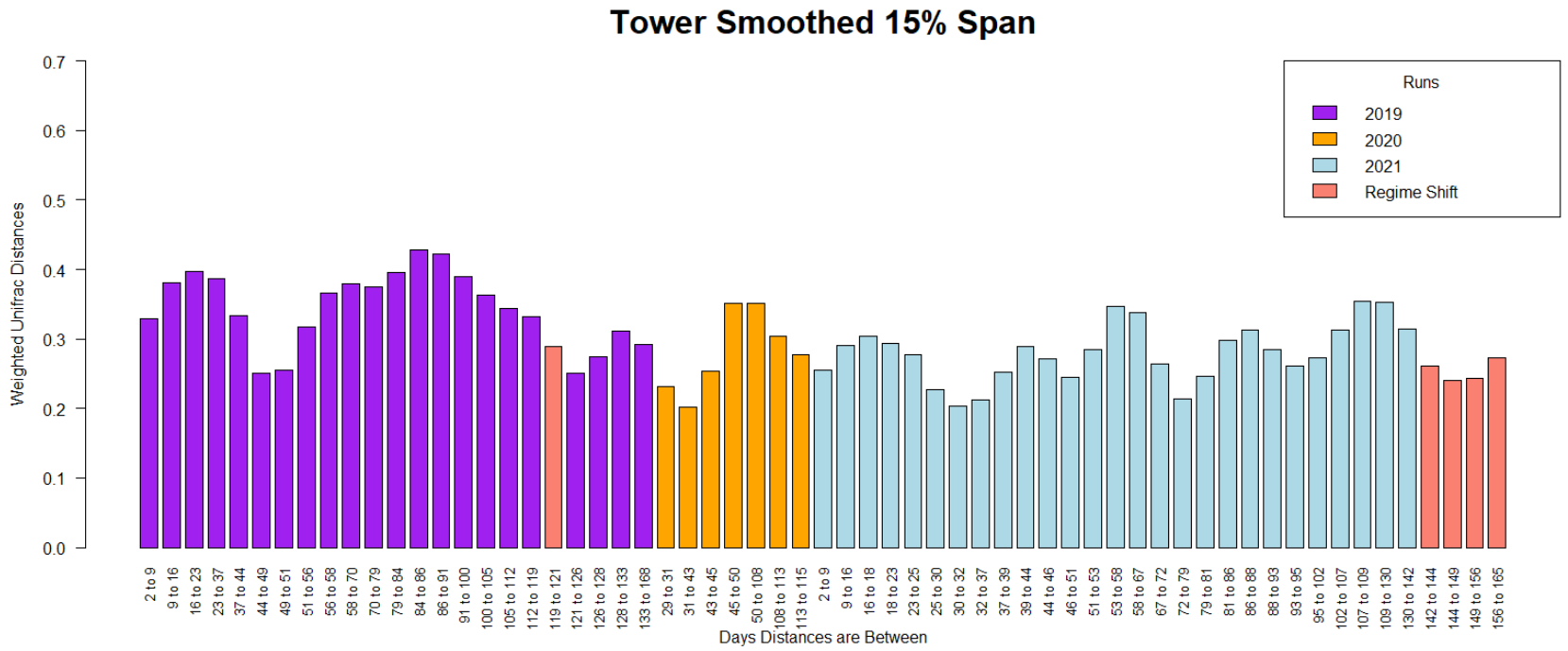

Figure 2.20 *Moving window analysis applied to successive weighted UniFrac distances for the T smoothed using LOESS*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). Regime shifts detected by the Student's t-test Analysis of Regime Shifts (STARS) are colored salmon. UniFrac distances were "smoothed" using Local Regression (LOESS). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

Results of the MWA for the LT are shown as barplots detailing the distances between successive samples (Figure 2.21), distances between the first timepoint in 2019 and all other samples (Figure 2.22), and successive distances smoothed using the LOESS technique (Figure 2.23).

In the LT, community composition between successive samples was, on average, at a distance of 0.33 ± 0.08 in 2019, 0.27 ± 0.19 in 2020, and 0.34 ± 0.14 in 2021 (Figure 2.21). In terms of percent similarity in microbial composition, successive samples in 2019 were on average 66.52 ± 8.40% similar, samples in 2020 were 73.18 ± 18.82 % similar, and samples in 2021 were 66.06 ± 13.75% similar. Using a two-sample, two-tailed t-test assuming unequal variance, average stability was not significantly different between all three runs in the LT (2019-2020: $p=0.45$; 2020-2021: $p=0.41$; 2019-2021: $p=0.91$). These results indicate stability, on average, is consistent over time in the LT.

For distances between successive samples (Figure 2.21), the lowest distance between successive samples (highest moment of stability) occurred in 2020 between days 17 and 22 (distance of 0.047, 95.33% similar) The highest distance between successive samples (highest moment of instability) occurred in 2021 between days 95 and 100 (distance of 0.58, 42.31% similar). In 2019, the LT microbiome was initially stable (below average distance for the run) then rose to a peak of instability by day 79. Stability generally increased throughout the rest of the run. In 2020, microbial stability was initially below average but increased between days 17 and 24 before a spike of instability occurred between days 24 and 29. Stability rose to about average for the rest of the run. In 2021, microbial stability rose during the beginning of the run until day 30. Stability was high between days 30 and 39. A spike in instability occurred after day 39 then

lowered again. This pattern of stability falling and rising occurred several times throughout the rest of the run.

For distances between the first timepoint in 2019 (Figure 2.22), the starting time point in the LT was day 16 in 2019. On average, stability in terms of the experimental start point was a distance of 0.34 ± 0.11 in 2019, 0.44 ± 0.08 in 2020, and 0.37 ± 0.13 in 2021 for an overall average of 0.37 ± 0.12. Using a two-sample, two-tailed t-test assuming unequal variance, average stabilities from the starting time in all three runs were not significantly different (2019-2020: $p=0.07$; 2020-2021: $p=0.08$; 2019-2021: $p=0.54$). This result suggests that, in viewing stability in terms of each sample to the experimental start, there was no significant difference in LT microbial stability across all three runs.

The lowest distance from the run start occurred in 2021 on day 39 (distance of 0.18) (Figure 2.22). The highest distance occurred in 2021 on day 100 (distance of 0.59). In 2019, distances were initially 0.23-0.29 then rose to 0.55 on day 79 before returning to 0.23 on the next time point. Distances rose to 0.41 by day 105 then fell for the rest of the run. The general trend of stability lowering on days 79 and 105 coincides with stability over time lowering between days 65 and 79 and days 98 to 105 (Figure 2.21). In 2020, distances initially rose from 0.31 on day 11 to 0.50 on day 17. Distances then lowered on day 29 to 0.34 before rising again to 0.45 by day 106. The period of lower stability between days 17 and 24 corresponds to the highest moments of stability over time in the LT as observed in Figure 2.21. Likewise, the rise in stability observed on day 29 corresponds to high instability between days 24 and 29 in Figure 2.21. In 2021, several peaks in distances of about 0.56 to 0.59 occurred on days 18, 23, 65, 72, 100, and 142.

Valleys (periods of high stability, shorter distances) of about 0.20 to 0.30 occurred between days 30 and 39, days 46 and 58, day 67, day 74, days 93 and 95, and days 102 and 109 with the lowest two distances being 0.18 on days 32 and 39. These peaks and valleys mostly coincide with those observed in 2021 in Figure 2.21.

The LOESS smoothing technique permits clearer analysis of regime shifts in timescale data. In Figure 2.23, the LT data was smoothed at a span of 25% to maximize the number of regime shifts detected. Figure 17c shows the distances between successive samples with the smoothing technique applied and regime shifts detected colored.

Smoothing stability data over time, displayed in Figure 2.23 showcased general trends in rising and falling stability already identified in Figure 2.21. In 2019, stability fell between days 16 and 84 before rising for the remainder of the run. In 2020, stability mostly rose throughout the run. In 2021, four main peaks occurred between days 11 and 30, 46 and 81, 93 and 107, and 142 onward. STARS detected regime shifts on days 4 to 11 and days 107 onward, all in 2021. The first shift preludes the rise in instability starting on days 11 to 16. The final detected shifts occur on the last nine time points in the experiment in the LT and includes a rise and fall in stability. The shifts detected from days 107 to 109, 109 to 114, and 114 to 130 may be indicators of the decreasing stability between days 130 and 151. The final detected shifts occurring on the last six time points in the experiment in the LT may suggest the observed decline in stability at the end of the experiment was significant.

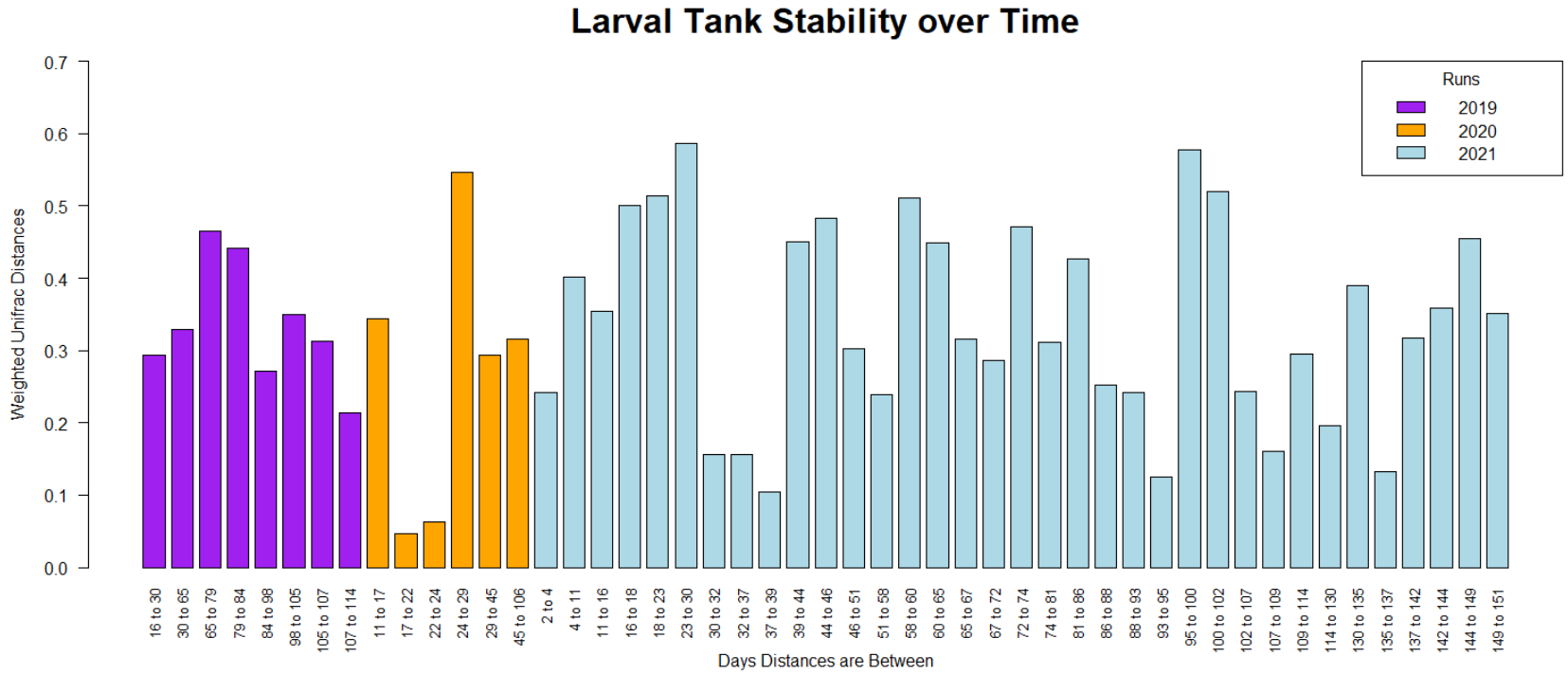

Figure 2.21 *Moving window analysis applied to successive weighted UniFrac distances for the LT*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

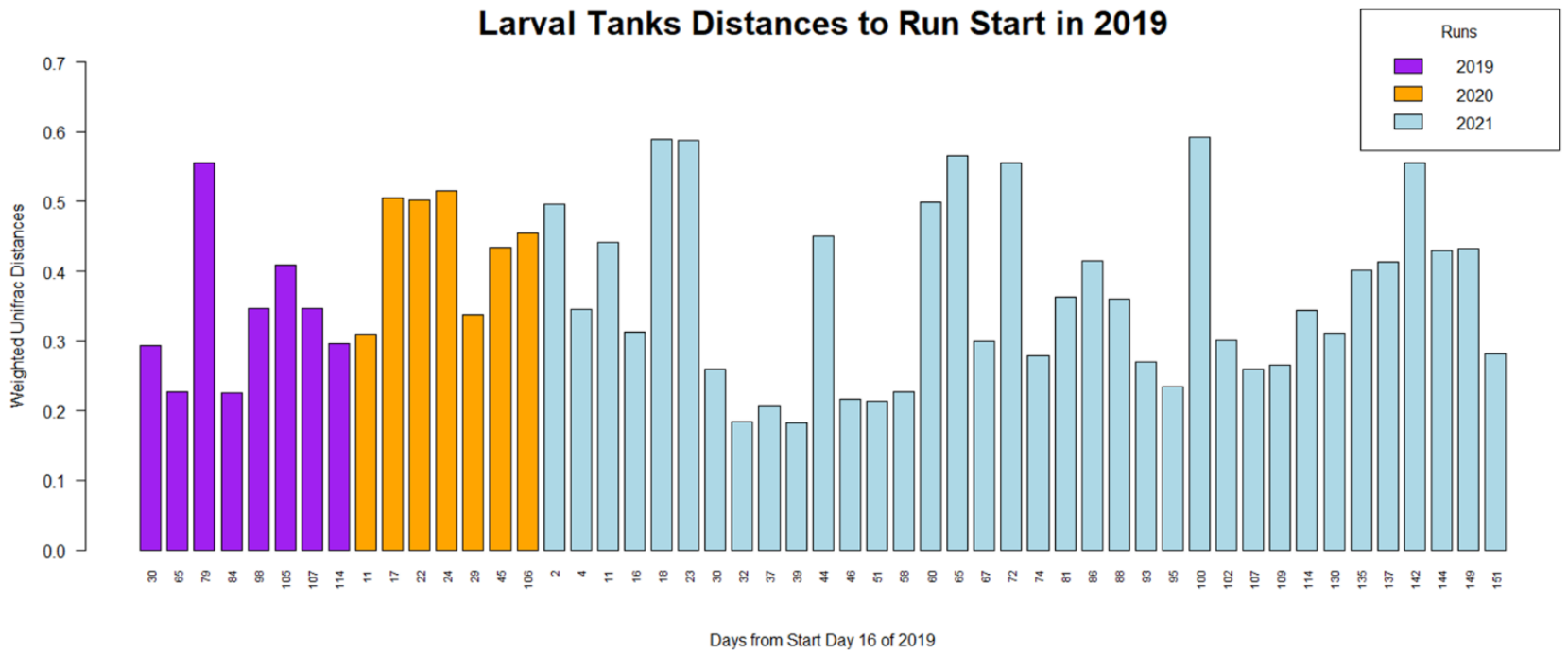

Figure 2.22 *Moving window analysis applied to the weighted UniFrac distances between the first sampling day and all subsequent samples in the LT*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between the beginning of the experiment (day 16 of 2019) and all subsequent samples (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

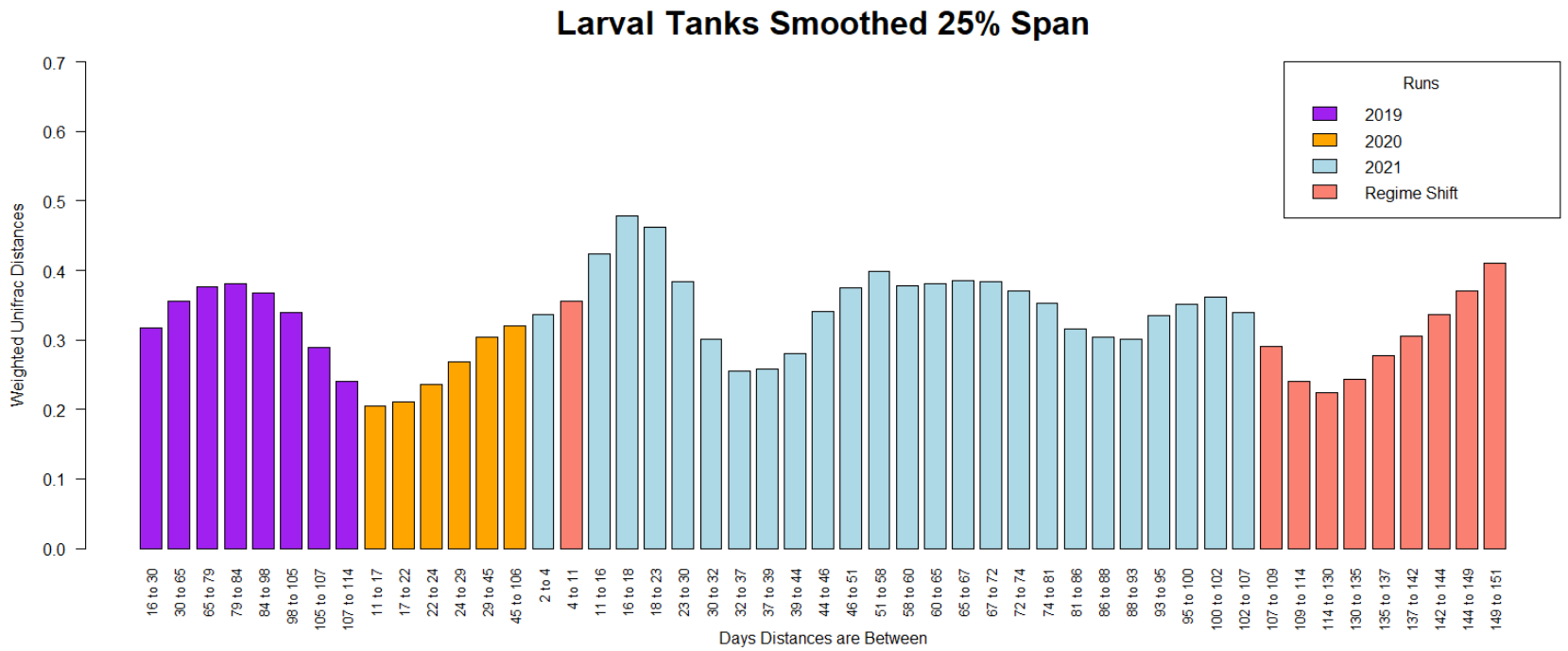

Figure 2.23 *Moving window analysis applied to successive weighted UniFrac distances for the LT smoothed using LOESS*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). Regime shifts detected by the Student's t-test Analysis of Regime Shifts (STARS) are colored salmon. UniFrac distances were "smoothed" using Local Regression (LOESS). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

Results of the MWA for the AH are shown as barplots detailing the distances between successive samples (Figure 2.24), distances between the first timepoint in 2019 and all other samples (Figure 2.25), and successive distances smoothed using the LOESS technique (Figure 2.26).

In the AH, community composition between successive samples was, on average, at a distance of 0.42 ± 0.18 in 2019, 0.41 ± 0.16 in 2020, and 0.34 ± 0.19 in 2021 (Figure 2.24). In terms of percent similarity, successive samples in 2019 were on average 57.62 ± 18.14% similar samples in 2020 were 59.49 ± 15.69% similar, and samples in 2021 were 66.39 ± 19.23% similar. Using a two-sample, two-tailed t-test assuming unequal variance, average stability was not significantly different between all three runs in the AH (2019-2020: $p=0.69$; 2020-2021: $p=0.06$; 2019-2021: $p>0.05$). These results indicate stability, on average, is consistent over time in the AH.

For distances between successive samples (Figure 2.24), the lowest distance between successive samples (highest moment of stability) occurred in 2021 between days 31 and 32 (distance of 0.02, 98.31% similar). The highest distance between successive samples (highest moment of instability) occurred in 2021 between days 44 and 45 (distance of 0.82, 18.47% similar). In all three runs, stability in the AH was sporadic with no discernable pattern. Even between AH1 (samples taken the day received) and AH2 (sample taken 24 hours later), there is no consistent stability or instability. This is likely due to the AH being the only compartment not within the RAS loop and the only compartment regularly replenished (new algae are received every other day). The AH has no means of maintaining a stable microbiome.

For distances between the first timepoint in 2019 (Figure 2.25), the starting time point in the AH was day 8 in 2019. On average, stability in terms of the experimental start point was a distance of 0.55 ± 0.19 in 2019, 0.51 ± 0.18 in 2020, and 0.50 ± 0.24 in 2021 for an overall average of 0.51 ± 0.22. Using a two-sample, two-tailed t-test assuming unequal variance, average stabilities from the starting time in all three runs were not significantly different (2019-2020: $p=0.45$; 2020-2021: $p=0.90$; 2019-2021: $p=0.37$). This result suggests that, in viewing stability in terms of each sample to the experimental start, there was no significant difference in RW microbial stability across all three runs.

The lowest distance from the run start occurred in 2021 on day 79 (distance of 0.06) (Figure 2.25). The highest distance occurred in 2021 on day 62 (distance of 0.86). As was observed in stability over time (Figure 2.24), stability from the starting time point in the AH was sporadic with no discernable pattern across all three runs. There are no matching of periods of stability or instability between the two views of stability in the AH.

The LOESS smoothing technique permits clearer analysis of regime shifts in timescale data. In Figure 2.26, the AH data was smoothed at a span of 20% to maximize the number of regime shifts detected. Figure 2.23 shows the distances between successive samples with the smoothing technique applied and regime shifts detected colored.

Smoothing stability data over time, displayed in Figure 2.26, showcased general trends in rising and falling stability otherwise undiscernible in Figure 2.24. In 2019, stability lowered from day 8 to day 78, after which stability increased until days 113 to 114 when a gradual decline occurred for the remainder of the run. In 2020, two peaks of

instability occurred between days 14 and 23 and days 50 and 100. In 2021, two peaks between days 15 and 24 and days 38 and 50 occurred before a period of steadiness that lasted until days 72 to 73. Stability then decreased until days 88 and 92. After day 92, stability gradually rose throughout the rest of the run. STARS detected regime shifts on days 2 to 8, 28 to 29, and 101 to 106 in 2020 and days 74 to 80, 106 to 107, and 150 onward in 2021. The first shift in 2020 preludes the first peak of instability observed in 2020. The next two shifts in 2020 follow the first and second peaks. In 2021, the first and second shifts detected bracket the highest peak of instability in the run. The final detected regime shifts occur on the last six time points in the experiment in the AH, suggesting the observed rise in stability at the end of the experiment was significant.

In comparing compartment stability over time, the T was the most stable across each run while the AH was generally the least stable. The expected trend of increased stability by the ends of each run was not observed in any compartment. However, a pattern appeared to emerge in the RW and LT in 2021. Cross examining the timeline of both compartments reveals peaks and drops in stability in the LT match or are slightly offset from those in the RW. The T also appears to have a “delayed reaction” with some increases in instability starting directly after a peak in the RW. These trends do not appear in the AH. Trends observed in stability over time in the RW and LT generally coincided with trends in stability viewed as the distance between the experimental starting point and all subsequent samples. The T saw few connections between the two, and the AH saw little to none.

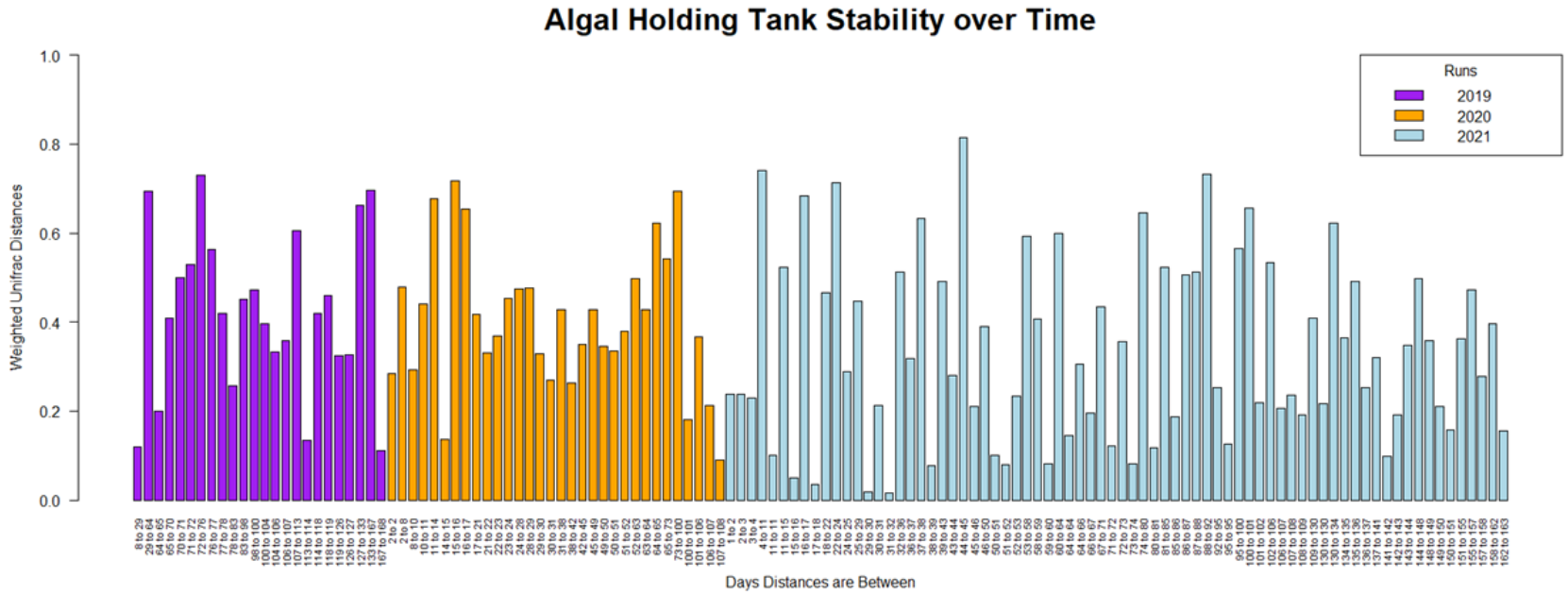

Figure 2.24 *Moving window analysis applied to successive weighted UniFrac distances for the AH*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

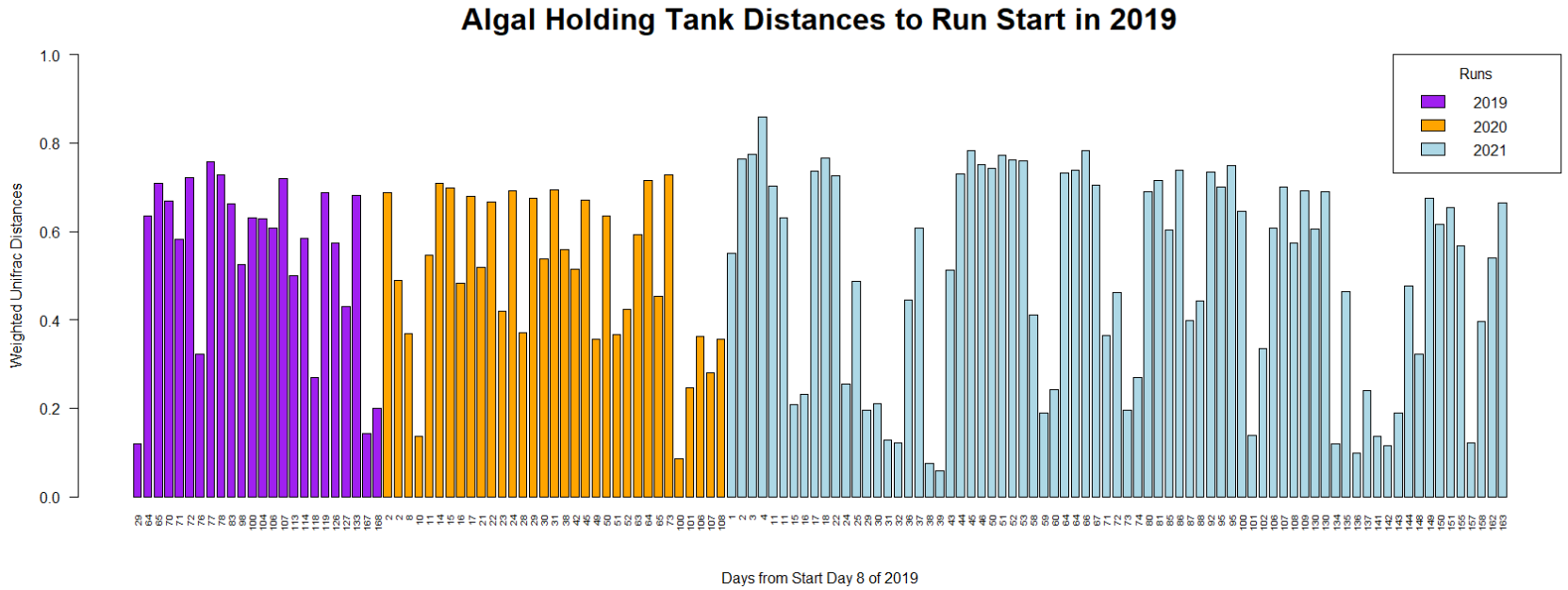

Figure 2.25 *Moving window analysis applied to the weighted UniFrac distances between the first sampling day and all subsequent samples in the AH*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between the beginning of the experiment (day 8 of 2019) and all subsequent samples (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently srt bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

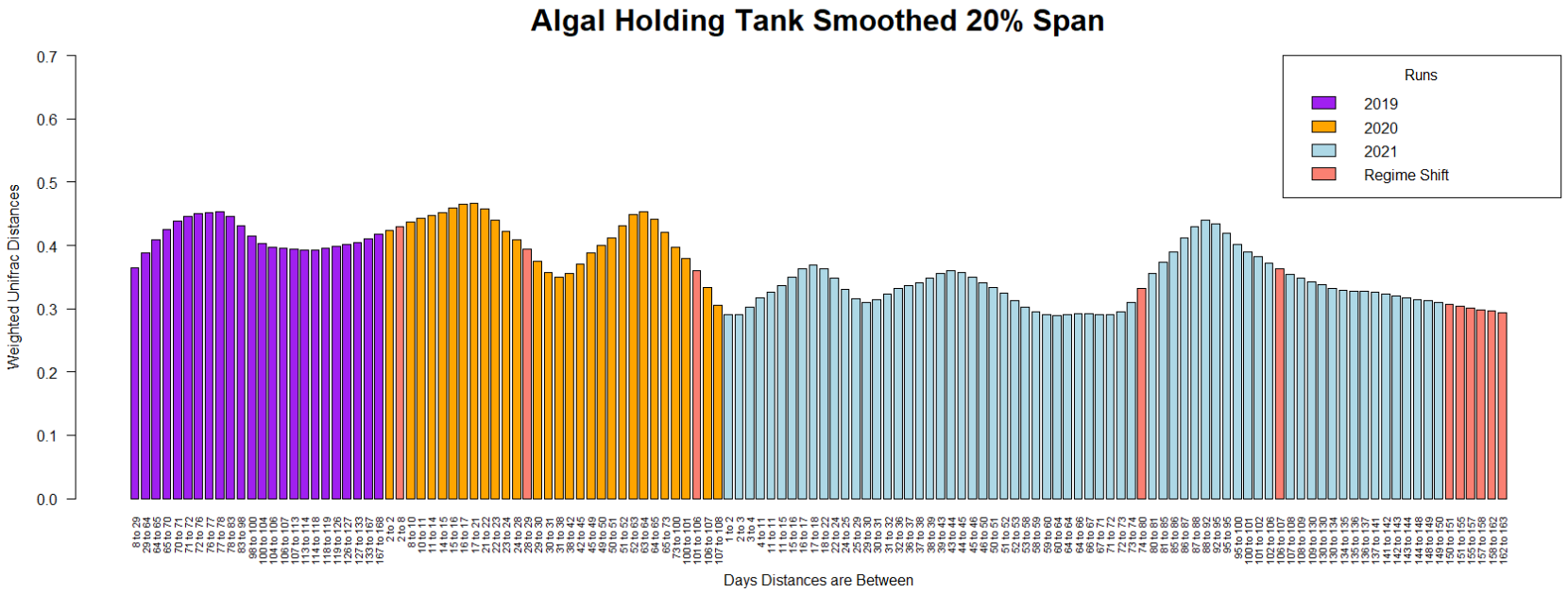

Figure 2.26 *Moving window analysis applied to successive weighted UniFrac distances for the AH smoothed using LOESS*

This analysis visualizes stability of microbiomes over time. UniFrac distance (y-axis) between successive days of sampling (x-axis) is displayed as bars spanning 2019 (purple), 2020 (orange), and 2021 (blue). Regime shifts detected by the Student's t-test Analysis of Regime Shifts (STARS) are colored salmon. UniFrac distances were "smoothed" using Local Regression (LOESS). The height of the bar can be used as an indicator of similarity between continuous time points. Consistently short bars between successive samples equates to high stability. Consistently high bars between successive samples equates to low stability.

Regime shifts were detected primarily in 2021 across all three runs with many shifts congregating in the last four to nine time points of the experiment. Though STARS did detect time points directly prior to rises and falls in stability, its reliability in detecting shifts in microbial stability is not fully understood. Further analysis of regime shifts in similar time series is recommended.

2.3.10 Larval Presence and Absence

A total of 22 broods were reared in the system in 2019, 4 were reared in 2020, and 11 were reared in 2021 (Table 15). In 2019, the average length of time a brood remained in the system was 10 days with a minimum time of four days (brood 8) and maximum of 20 days (brood 1). In 2020, the average length of time a brood was reared in the system was 14 days with a minimum time of 6 days (brood 1) and maximum of 17 days (brood 3). In 2021, the average length of time a brood remained in the system was 13 days with a minimum of 2 days (brood 9) and maximum of 21 days (brood 8). Most successful broods were harvested over the course of one to five days. Unsuccessful broods varied in rearing length depending on mortality rates. In 2019 and 2021, broods sometimes overlapped since four LTs were available, and it was often more efficient for production to rear more than one brood in the system at a time. In 2020, fewer broods were spawned and reared due to COVID19 restrictions.

Table 2.15 *: Days broods were added and removed from the larval tanks in each run*

Days in 2019	Broods added/removed	Days in 2020	Brood added/removed	Days in 2021	Brood added/removed
1	2019_1-added	9	2020_1-added	3	2021_1-removed
15	2019_2-added	15	2020_1-removed	4	2021_1-removed
21	2019_1-removed	16	2020_2-added	5	2021_1-removed
24	2019_3-added	29	2020_2-removed	6	2021_1-removed
25	2019_2-removed	30	2020_2-removed	7	2021_1-removed
29	2019_4-added	31	2020_2-removed	11	2021_2-added
36	2019_5-added	37	2020_3-added	24	2021_2-removed
37	2019_4-removed	50	2020_3-removed	25	2021_2-removed
38	2019_3-removed	51	2020_3-removed	26	2021_2-removed 3-added
39	2019_3-removed	52	2020_3-removed	27	2021_4-added
40	2019_3-removed	53	2020_3-removed	31	2021_4-removed
41	2019_3-removed	54	2020_3-removed	38	2021_5-added
42	2019_3-removed	58	2020_4-added	44	2021_3-removed
45	2019_5-removed	72	2020_4-removed	51	2021_5-removed
50	2019_6-added	73	2020_4-removed	52	2021_5-removed
57	2019_6-removed	74	2020_4-removed	67	2021_6-added
64	2019_7-added	75	2020_4-removed	73	2021_6-removed
71	2019_7-removed			80	2021_7-added
74	2019_8-added			95	2021_7-removed 8-added
78	2019_8-removed 9-added			96	2021_7-removed
81	2019_10-added			97	2021_7-removed
84	2019_9-removed			113	2021_8-removed
85	2019_11-added			116	2021_8-removed
87	2019_10-removed			127	2021_9-added
92	2019_12-added			129	2021_9-removed 10-added
94	2019_13-added			145	2021_10-removed
99	2019_14-added			146	2021_10-removed 11-added
100	2019_13-removed			163	2021_11-removed
103	2019_11-removed				
106	2019_15-added				
107	2019_14-removed				
110	2019_12- 15-removed				
111	2019_12-removed				
112	2019_16-added				
115	2019_17-added				
122	2019_16-removed				

Table 2.15 (continued).

115	2019_17-added				
122	2019_16-removed 18-added				
123	2019_17-removed				
126	2019_19-added				
129	2019_20-added				
131	2019_18-removed				
135	2019_20-removed				
137	2019_21-added				
139	2019_19-removed				
149	2019_21-removed				
150	2019_21-removed 22-added				
163	2019_22-removed				
164	2019_22-removed				
165	2019_22-removed				
166	2019_22-removed				

An Analysis of Similarity (ANOSIM), which ranked similarities in data using the Spearman rank correlation, was used to compare samples taken when larvae were present in the RAS to samples taken when larvae were absent across all three runs in each within-loop compartment (Table 16). The calculated R statistic, which falls between +1 and -1, informs whether the microbial composition of samples in one group is similar to the composition of samples in the other group. The closer to 0 R is, the higher the chance the null hypothesis of no difference is accepted. A significance level of less than 0.05 determines whether (dis)similarities detected by the R statistic are significant.

The ANOSIM results indicate the RW microbiome when larvae were present and absent was significantly different only in 2019. In 2020 and 2021, the RW microbiome was similar regardless of larval presence or absence. Similarly, the T microbiome differed significantly when larvae were present or absent only in 2019. In 2020 and 2021, the T microbiome was similar regardless of larval presence or absence. In the LT, an

insufficient number of samples taken when larvae were absent in 2019 resulted in no test being performed. In 2020 and 2021, the LT microbiome was similar in the presence and absence of larvae.

Table 2.16 *One-way analysis of similarity (ANOSIM) of larval presence or absence*

The analysis was performed to determine the similarity in microbial composition between samples when larvae were present and absent in all within-loop compartments across all three years. Bolded text indicates a significant result ($p<0.05$).

One-Way ANOSIM (Analysis of Similarities) of Larval Presence and Absence					
Factors					
Place	Name	Type	Levels		
A	Larval Pres/Abs	Unordered	2		
Compartment	Run	Sample statistic (R)	Significance level	Perms	Perms ≥ R
Raceway	2019	0.57	**0.014**	276	4
	2020	-0.103	0.776	495	384
	2021	-0.053	0.636	9999	6330
Tower	2019	0.421	**0.035**	2024	70
	2020	0.231	0.161	56	9
	2021	0.243	0.108	9999	1078
Larval Tanks	2019	NA	NA	NA	NA
	2020	0.067	0.286	7	2
	2021	-0.006	0.46	9999	4603

In summary, both the RW and T microbiome differed in the presence and absence of larvae in 2019 but not in 2020 or 2021. No test was performed in 2019 in the LT, but a similar result of no difference was found in 2020 and 2021. Further analysis of the relationship between larvae reared in the system and the system's microbiomes are explored in chapter 3.

2.3.11 Experimental Confounders in 2019

During the 2019 run, a broodstock holding system was added to the RAS from days 25 to 90. This addition was made in response to the second opening of the Bonnet

Carre Spillway during May and June of 2019. The opening of the spillway presented an immediate hazard to the survival of living oysters in the Mississippi Sound, including the broodstock adults TCMAC maintained in the Sound. Encroaching freshwater from the spillway into the Mississippi Sound threatened the livelihood of the broodstock, so the surviving adult oysters were recovered and brought to the RAS facility in June. Approximately 1562 broodstock (BS) were salvaged and added to an available tank and connected to the RAS (Figure 2.27). This addition required 9000 L of artificial seawater to be made in the Raceway and subsequently circulated through the RAS. During June and July, the BS and larvae reared in the RAS (broods 3 through 11) were monitored for mortalities. The BS remained connected to the RAS for 65 days during which approximately 803 BS died and broods 5 through 10 were terminated for low larval health. TCMAC staff suspected copper contamination was the cause due to a prior experiment involving the tank the BS were added to. The BS tank was removed from the RAS on day 90. To combat RAS larval mortality, TCMAC staff added CupriSorb, a copper/heavy metals absorbent, to the T between August (day 102) and December (post-sampling). More EDTA was added all three stages of water reclamation on day 107.

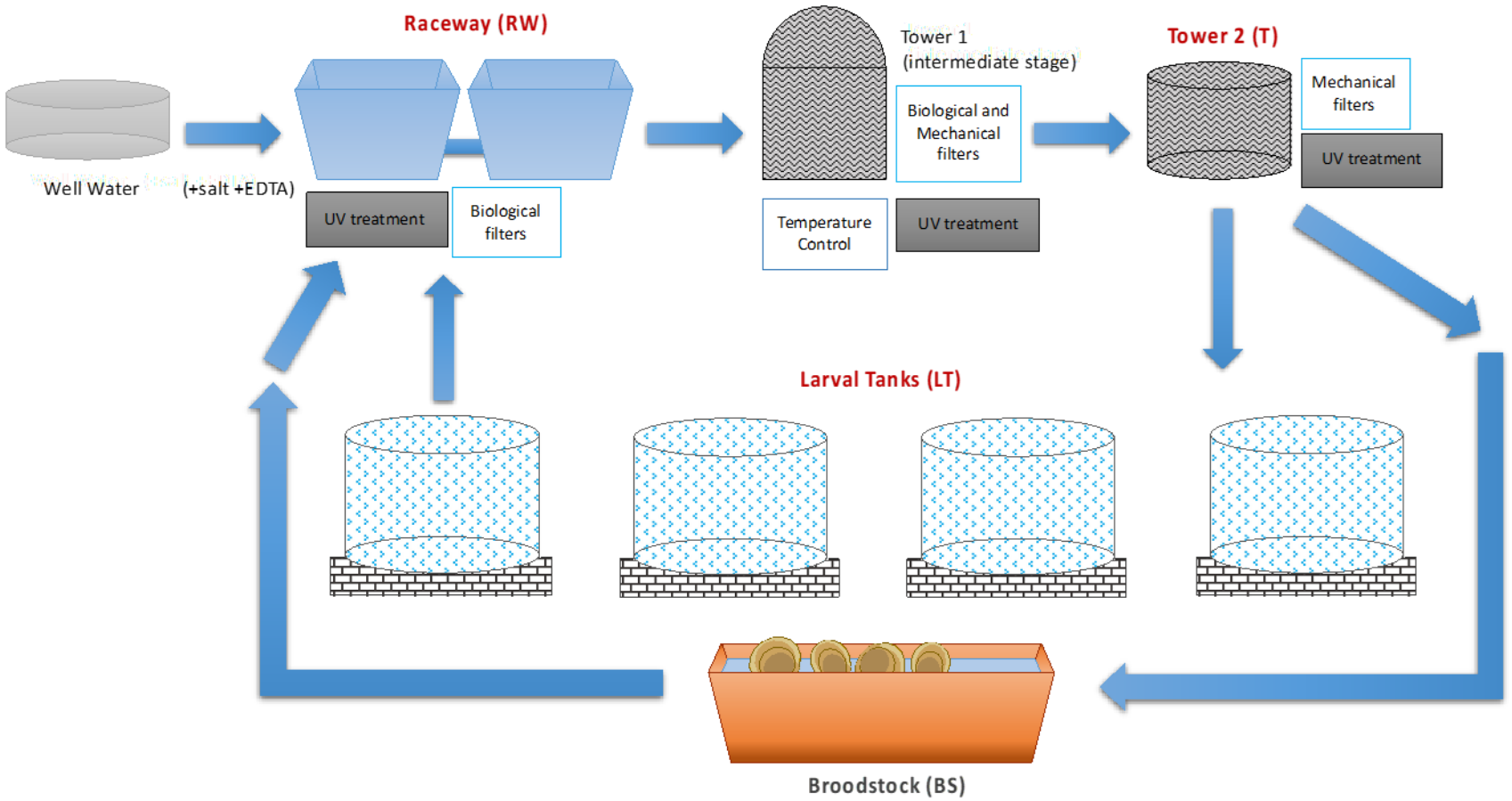

Figure 2.27 *Diagram depicting the system with the broodstock (BS) tank added between days 25 and 90 of 2019*

A trace metals analysis performed on day 147 is summarized in Table 2.17. In the RW, copper concentrations were about 0.75 ppb, lead concentrations were about 0.02 ppb, and zinc concentrations were about 13 ppb. In the BS, copper concentrations were about 30 ppb, lead concentrations were about 0.08 ppb, and zinc concentrations were about 49 ppb. Concentrations in the RW are not lethal to *C. virginica* larvae, but concentrations in the BS are (Calabrese *et al.*, 1973, Weng & Wang, 2014). It is noted the trace metals analysis was performed 57 days after the BS was removed and 45 days CupriSorb treatment began in the RAS. This analysis proves copper, zinc, and lead all contaminated the RW and BS. Coupled with larval mortality taking place at the same time, it can be said the larvae and the water in the RAS were contaminated but to an unknown degree.

Table 2.17 *Trace metals analysis performed on day 147 of 2019*

Concentrations of copper, lead, and zinc in the RAS's Raceway (RW) and in the broodstock tank (BS). Analysis performed after 45 days of CupriSorb treatment in the RAS. Three replicate samples were collected per compartment.

Trace Metals Analysis				
Compartment	Replicate	Cu (ppb)	Pb (ppb)	Zn (ppb)
RW	1	0.75	0.031	13.1
	2	0.76	0.012	12.3
	3	0.74	0.015	12.2
	Average	0.75	0.019	12.5
BS	1	29.94	0.079	48.6
	2	29.84	0.081	48.7
	3	30.25	0.070	48.8
	Average	30.01	0.077	48.7

After sampling in 2019 had ended, specifically from December 2nd to 16th, approximately 8700 gallons of new water was made up in the RAS to ensure the contamination was removed from the system. This amounted to over 90% of the system's

water volume being remade, effectively restarting the system. The BS, contamination, and treatments are confounding variables that had the potential to impact the microbiomes being studied in this work. This, the specific effects from disturbances occurring in 2019 were examined. To facilitate this examination, samples from 2019 were subdivided into four groups, corresponding to when the system was operated under standard procedures (2019a), when the BS was added (2019b), when the BS was removed (2019c), and when treatments had been administered (2019d) as visualized in Figure 2.28.

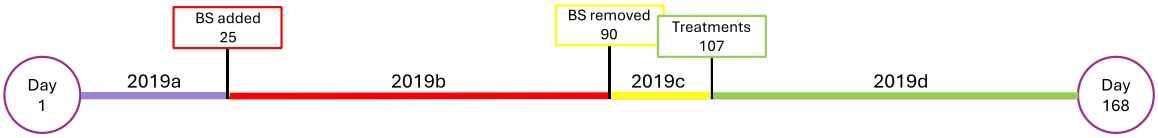

Figure 2.28 *Timeline of events in 2019 detailing the four subdivisions of 2019*

Subdivisions are as follows: standard system operations (2019a), BS addition (2019b), BS removal (2019c), and administration of treatments (2019d).

The RW dataset was revisualized via an nMDS plot in Figure 2.29 with the added subdivisions in 2019. Samples in 2019a are included in a cluster with 2020 and 2021 samples but are not close together. Most samples in 2019b make up a tight cluster in the bottom-center of the two-dimensional graph. Additionally, one of the outlier samples belongs to 2019b. Samples from 2019c are spread across the lower half of the plot, making up clusters with 2019b and 2019d samples. 2019d samples are also spread across the lower half of the plot. Some form clusters with 2019b and 2019c samples while others are spread out away from any other RW samples. Most notable is the clear grouping of samples taken while the BS was connected directly to the RW. The two samples taken

before the BS was added group with 2020 and 2021 samples. Samples taken after the BS removal do not form groups with samples taken in 2020 or 2021.

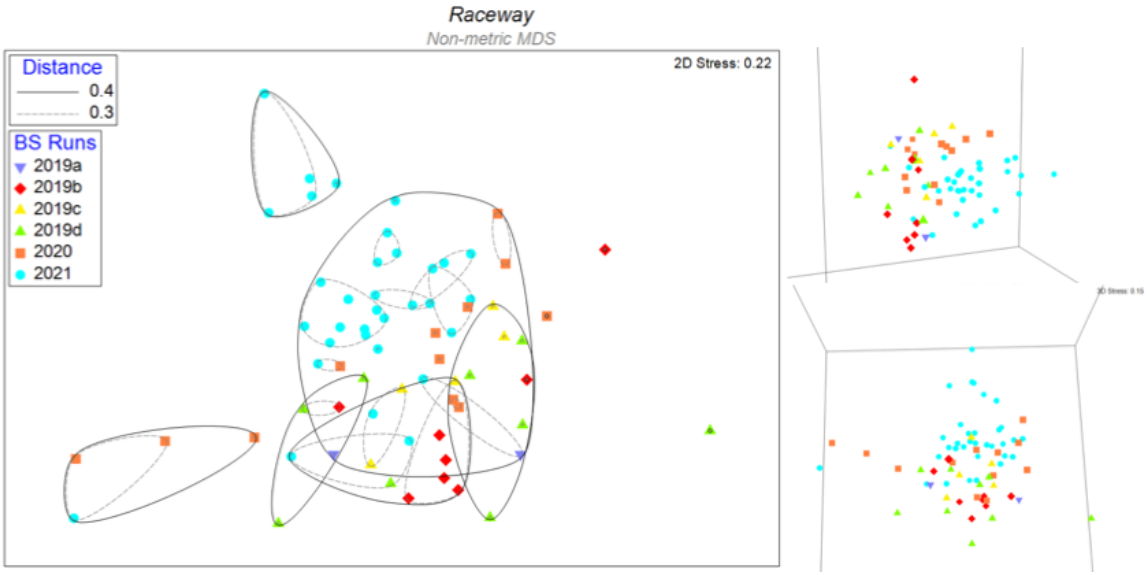

Figure 2.29 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the RW during all three runs, including 2019 subdivisions, in two- and three-dimensional space*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by run, included a subdivided 2019, for visualization. Samples are grouped by runs colored orange for 2020 and blue for 2021. Modified 2019 sections are 2019a (purple) pre-broodstock, 2019b (red) during broodstock addition, 2019c (yellow) post-broodstock, and 2019d post-Cu treatments. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).

To determine specific differences between samples in the four subdivisions in 2019 and samples in 2020 and 2021, a PERMANOVA pairwise test was performed on the RW data included in Figure 2.29, as shown in Table 2.18. The pairwise test shows samples taken after confounders were introduced were generally different from samples in 2020 and 2021. 2019a was not different from any other 2019 subdivision nor 2020 or 2021. The pair with the highest t test score was 2019b and 2019d, indicating the

community present when contamination was ongoing was distinct from the community present during treatment administrations.

Table 2.18 *Permutational multivariate analysis of variance (PERMANOVA) pair wise tests were performed on the RW for subdivided run 2019, 2020, and 2021 to determine differences in the microbiome between pairs*

Bolded text indicates a significant result ($p<0.05$).

Pair Wise Tests PERMANOVA on the RW			
Sums of squares type: Type III (partial)			
Fixed effects sum to zero for mixed terms			
Permutation method: Unrestricted permutation of raw data			
Number of permutations: 9999			
Factors			
Name	Type	Levels	
BS Runs	Fixed	6	
Groups	t	P(perm)	Unique perms
2019d, 2019a	1.211	0.0868	55
2019d, 2019b	1.5244	**0.009**	8102
2019d, 2019c	1.1539	0.1307	1992
2019d, 2021	2.1883	**0.0001**	9912
2019d, 2020	1.4549	**0.022**	9752
2019a, 2019b	1.2194	0.1323	45
2019a, 2019c	1.1692	0.1438	21
2019a, 2021	1.3497	0.0536	595
2019a, 2020	1.0907	0.233	91
2019b, 2019c	1.392	0.0581	1285
2019b, 2021	2.0966	**0.0001**	9909
2019b, 2020	1.6849	**0.0041**	9543
2019c, 2021	1.673	**0.0023**	9799
2019c, 2020	1.1081	0.2381	4912
2021, 2020	1.7024	**0.0023**	9923

To further delve into factors shaping community composition in the RW with a focus on confounders in 2019, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels. The algorithm determined which RW samples belonged to one of ten clusters (Figure 2.30). Cluster A contains mostly 2019c, 2019d, and 2020 samples, cluster B contains 2019b, 2019d, and 2020

samples, cluster E contains mostly 2021 samples and one of the two 2019a samples, cluster G consists primarily of 2019b samples, and cluster H includes 2019a, 2019c, 2020, and 2021 samples. Clusters C, D, F, I, and J include primarily 2020 and 2021 samples.

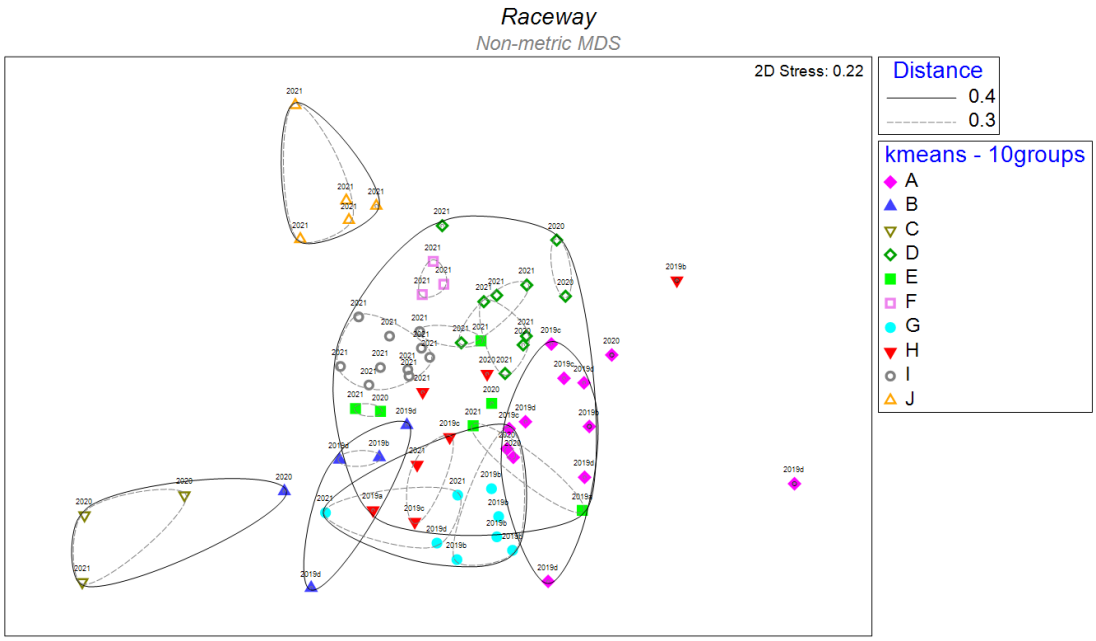

Figure 2.30 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the RW during all three runs and organized into k-means clusters*

. Data derived from the weighted UniFrac distance matrix. Samples are grouped by ten k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).

The T dataset was revisualized via an nMDS plot in Figure 2.31 with the added subdivisions in 2019. Samples in 2019a are spread across the right side of the plot alongside other 2019 samples. Most samples in 2019b make up a tight cluster at a distance of 0.35 (65% similarity). Samples from 2019c occupy the right side of the plot near samples from 2019b. 2019d samples are mostly grouped in the center of the plot

above samples from 2019b. Visually, the distinction between 2019 and the latter two runs is evident as samples from 2019 primarily spread across the right side of the plot while samples from 2020 and 2021 are more tightly grouped in their respective runs on the left side of the plot.

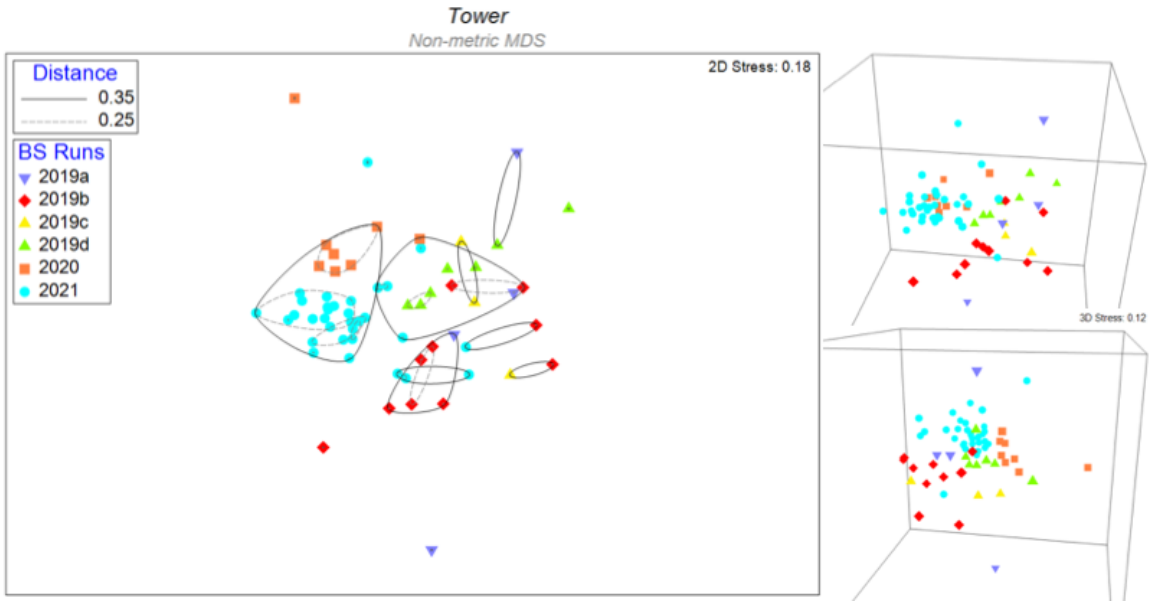

Figure 2.31 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the T during all three runs, including 2019 subdivisions, in two- and three-dimensional space*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by run, included a subdivided 2019, for visualization. Samples are grouped by runs colored orange for 2020 and blue for 2021. Modified 2019 sections are 2019a (purple) pre-broodstock, 2019b (red) during broodstock addition, 2019c (yellow) post-broodstock, and 2019d post-Cu treatments. A CLUSTER analysis generated distance contour lines at 0.35 (65% similar) and 0.25 (75% similar).

To determine specific differences between samples in the four subdivisions in 2019 and samples in 2020 and 2021, a PERMANOVA pairwise test was performed on the T data included in Figure 2.31, as shown in Table 2.19. The pairwise test shows samples taken from 2019c were not significantly different from the other samples taken from 2019. All other pairs were significantly different from each other. The pair with the

highest t test score was 2020 and 2021, suggesting the difference between runs overall is greater than the differences between subdivisions in 2019. These results indicate the BS event had an impact on the T microbiome, but the event was no more impactful than the effect of time in this compartment.

Table 2.19 *Permutational multivariate analysis of variance (PERMANOVA) pair wise tests were performed on the T for subdivided run 2019, 2020, and 2021 to determine differences in the microbiome between pairs*

Bolded text indicates a significant result ($p<0.05$).

Pair Wise Tests PERMANOVA on the Tower			
Sums of squares type: Type III (partial)			
Fixed effects sum to zero for mixed terms			
Permutation method: Unrestricted permutation of raw data			
Number of permutations: 9999			
Factors			
Name	Type	Levels	
BS Runs	Fixed	6	
Groups	t	P(perm)	Unique perms
2019d, 2019a	1.322	**0.0181**	330
2019d, 2019b	1.4853	**0.01**	7744
2019d, 2019c	1.2293	0.0846	120
2019d, 2021	2.5426	**0.0001**	9895
2019d, 2020	2.2961	**0.0006**	5073
2019a, 2019b	1.3323	**0.0329**	1000
2019a, 2019c	1.2602	0.0574	35
2019a, 2021	2.3515	**0.0001**	9118
2019a, 2020	2.2796	**0.0016**	495
2019b, 2019c	1.2595	0.0924	286
2019b, 2021	2.7499	**0.0001**	9916
2019b, 2020	2.68	**0.0001**	8882
2019c, 2021	2.0038	**0.0002**	5081
2019c, 2020	2.079	**0.0062**	165
2021, 2020	2.5876	**0.0001**	9921

As was done for the RW, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels in the T. The algorithm organized samples into ten clusters and the results were visualized in an nMDS plot in

Figure 2.32. Cluster A contains samples from all subdivisions as well as 2020 and 2021, cluster B is comprised entirely of 2019b samples, cluster C includes samples from 2019b, 2019c, 2019d, 2020, and 2021, cluster D is one 2019d sample, cluster G is one samples from 2019a, cluster H includes mostly 2019b and 2021 samples, and cluster I is one sample from 2019a. Groups E, F, and J are comprised entirely of 2020 or 2021 samples.

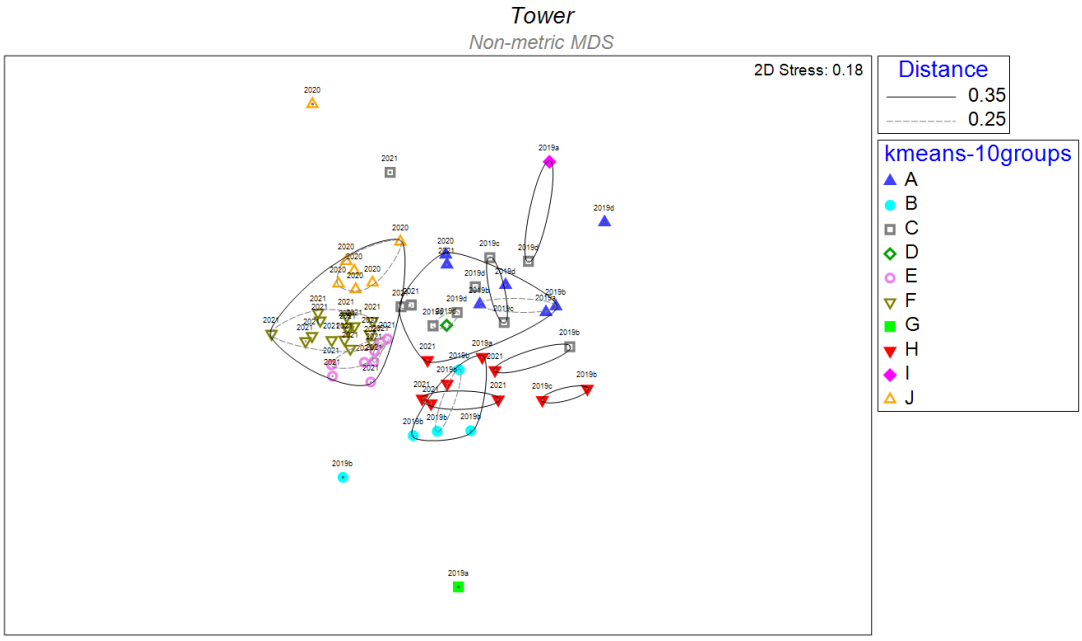

Figure 2.32 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the T during all three runs and organized into k-means clusters*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by ten k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.35 (65% similar) and 0.25 (75% similar).

The LT dataset was revisualized via an nMDS plot in Figure 2.33 with the added subdivisions in 2019. The single sample from 2019a is oriented near samples from 2021. 2019b samples are spread across the plot near 2021 samples. 2019c samples are oriented mostly with 2019b and 2020 samples. The single sample from 2019d is located in the

center of the plot near 2021 samples. There does not appear to be a pattern of orientation by 2019 subdivisions.

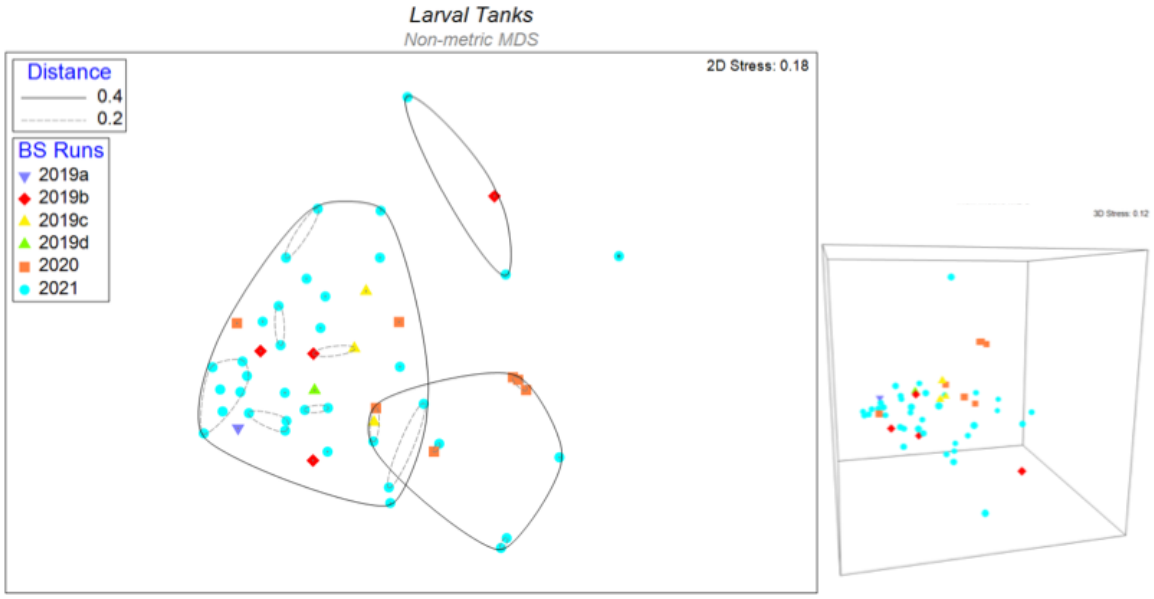

Figure 2.33 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the LT during all three runs, including 2019 subdivisions, in two- and three-dimensional space*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by run, included a subdivided 2019, for visualization. Samples are grouped by runs colored orange for 2020 and blue for 2021. Modified 2019 sections are 2019a (purple) pre-broodstock, 2019b (red) during broodstock addition, 2019c (yellow) post-broodstock, and 2019d post-Cu treatments. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

To determine specific differences between samples in the four subdivisions in 2019 and samples in 2020 and 2021, a PERMANOVA pairwise test was performed on the LT data included in Figure 2.33, as shown in Table 2.20. The pairwise test shows only 2019b and 2020 as well as 2020 and 2021 were significantly different from each other. The pair with the highest t test score was 2020 and 2021. These results confirm the lack of grouping by 2019 subdivisions visualized in Figure 2.33, indicating there are generally no differences in the LT microbiome throughout the BS event in 2019.

Table 2.20 *Permutational multivariate analysis of variance (PERMANOVA) pair wise tests were performed on the LT for subdivided run 2019, 2020, and 2021 to determine differences in the microbiome between pairs*

Bolded text indicates a significant result ($p<0.05$).

Pair Wise Tests PERMANOVA on the Larval Tank			
Sums of squares type: Type III (partial)			
Fixed effects sum to zero for mixed terms			
Permutation method: Unrestricted permutation of raw data			
Number of permutations: 9999			
Factors			
Name	Type	Levels	
BS Runs	Fixed	6	
Groups	t	P(perm)	Unique perms
2019a, 2019b	0.84677	0.5942	5
2019a, 2019c	1.2923	0.2461	4
2019a, 2019d	No test		
2019a, 2021	0.87741	0.533	37
2019a, 2020	1.3619	0.2513	8
2019b, 2019c	0.9519	0.5278	35
2019b, 2019d	0.79231	0.5972	5
2019b, 2021	1.0616	0.3202	9415
2019b, 2020	1.6167	**0.0415**	330
2019c, 2019d	0.70212	1	4
2019c, 2021	0.97893	0.4254	6073
2019c, 2020	1.0396	0.3558	120
2019d, 2021	0.71298	0.8922	37
2019d, 2020	0.96717	0.3785	8
2021, 2020	1.7041	**0.0128**	9932

As was done for the RW and T, a k-means cluster algorithm was used to visualize the inherent grouping of samples independent of factor labels in the LT. The algorithm organized samples into ten clusters and the results were visualized in an nMDS plot in Figure 2.34. All clusters save G are comprised mostly of 2020 and/or 2021 samples. Cluster G is comprised of mostly 2019b samples, which reflects the PERMANOVA results singling out 2019b as the only differentiated 2019 subdivision (Table 2.20).

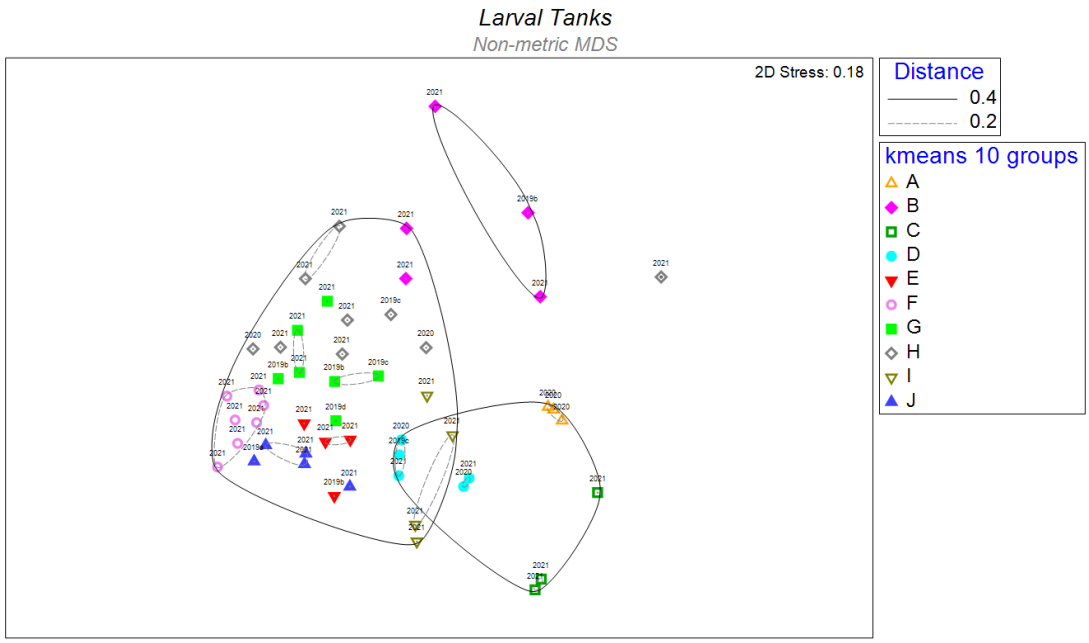

Figure 2.34 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the LT during all three runs and organized into k-means clusters*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by ten k-means clusters for visualization. Numbers above each symbol indicate the run the sample was taken. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar).

The events of 2019 did have an evident impact on the RW and T microbiomes in addition to the decline in larval health observed at the time. When the system was changed with the addition of the BS, signature communities emerged in the RW and T that were distinct from other periods in 2019, 2020, and 2021. A common trend of 2019 being singled out as differentiated from 2020 and 2021 across analyses of both compartments can be explained by the complete system reset between the end of sampling in 2019 and the start of sampling in 2020.

2.4 Discussion

2.4.1 Space and Time Influence the RAS Microbiome

Few studies have investigated microbial stability in a RAS (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Bakke *et al.*, 2017, Dahle *et al.*, 2022) and even fewer have focused on RASs rearing oyster larvae (Asmani *et al.*, 2016). This study is the first to assess microbial stability in an RAS rearing *C. virginica* larvae over a multi-year period.

Previous studies have determined the RAS microbiome is defined by the initial source water (Bartelme *et al.*, 2019), physicochemical conditions (Blancheton *et al.*, 2013, Bakke *et al.*, 2017, Rud *et al.*, 2017, Almeida *et al.*, 2021), and system design (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Asmani *et al.*, 2016, Almeida *et al.*, 2021). Here, water was consistently sourced from a well on-site and made into artificial seawater, keeping this factor constant throughout all three runs. Physicochemical conditions generally became steadier over time in both the RW and T (Figure 2.4), and the reduced variability in water quality parameters was reflected in lower spearman rank correlation coefficients in 2021 for both compartments (Tables 2.9,2.11). With source water held constant and physicochemical conditions becoming steadier over time, it was observed the RAS microbiome was impacted by the length of time the system was in operation and the system design.

It was hypothesized selective pressure on the RAS microbiome would result in no differences between the microbiome by the ends of each run across the system compartments. A lack of differentiation between the ends of each run was not observed. Although the day a sample was taken held no influence over community composition in

the RAS, the run in which a sample was taken did drive changes in community composition across all compartments.

The RAS microbiome was dominated by classes Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia with some variation between compartments observed at the family level (Figure 2.5). Beginning with the rearing tanks (LT) and first stage of water reclamation (RW), the RW and LT were comprised of Rhodobacteraceae, Alteromonadaceae, Flavobacteriaceae, and Pseudomonadaceae, but relative abundances of each were generally higher in the LT. The two were further differentiated by the RW's higher abundance of Cellvibrionaceae and the LT's higher abundance of Thalassospiraceae. That both the rearing tank (LT) and the first stage of water reclamation (RW) were comprised mostly of Proteobacteria is consistent with prior characterizations of the water microbiome of hatcheries (Stevick *et al.*, 2019, Arfken *et al.*, 2021) and RASs (Asmani *et al.*, 2016) rearing oyster larvae. Declining abundances of Alphaproteobacteria over time seen in the RW were also consistent with trends found in the RASs Asmani, *et al.* (2016) studied.

The community composition observed in the LT consisted of families previously associated as part of a larval core microbiome (as defined by Arfken *et al.*, 2021) and RASs rearing *C. virginica* and *C. gigas* larvae (Arfken *et al.*, 2021, Asmani *et al.*, 2016, Laroche *et al.*, 2018). Specifically, the family Rhodobacteraceae had the highest relative abundance (22.8%) in the larval core microbiome (Arfken *et al.*, 2021) and in the microbiome of the LT (overall avereage 36.7%) compared to the RW (overall average of 17.8%), T (overall average of 18.7%), and AH (overall average of 6.0%) microbiomes (Figure 2.5). Families Alteromonadaceae and Flavobacteriaceae were also prevalent in

the larval core microbiome (2.5% and 9.5%, respectively) (Arfken *et al.*, 2021) and had an increased presence in the LT microbiome (overall averages of 17.5% and 7.0%, respectively) compared to the other in-loop compartments (overall averages of 10.9% and 6.0% in RW and <5% in T) (Figure 2.5). Though no source analysis was performed in this study, taxa associated with oyster larvae being prevelant in the LT suggests the larvae were a source for the LT microbiomes. It is likely, given the direction of water flow from the LTs directly to the RW and the shared taxa between compartments, the LT microbiome was a source for the RW microbiome. Differences between the RW and LT microbiomes may be attributed to differences in compartment design and role in the RAS. As the first stage of water reclaimation, new water was made up only in the RW, and all water entering the RW was filtered through biological filters and UV treatment. Similarities between the larval core microbiome and the RAS microbiomes may reflect the importance of the water microbiome to the development of the larval microbiome (Asmani *et al.*, 2016, Laroche *et al.*, 2018) or the influence the larvae have to sourcing the RAS microbiome. More research is encouraged to determine the direction of influence.

The T microbiome consisted of the same dominant classes (Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia) in generally lesser relative abundances compared to the RW and LT (62.7% in T, 82.7% in RW, and 95.8% in LT, cumulatively). Like the in RW and LT, Rhodobacteraceae had the highest relative abundance in the T, though its abundance was usually lower in the T microbiome (averages of 16.3% in 2019, 32.8% in 2020, and 6.9% in 2021) than the RW and LT microbiomes (RW: averages of 33.3% in 2019, 9.2% in 2020, and 10.9% in 2021) (LT: averages of 49.9% in 2019, 22.8% in 2020,

and 37.5% in 2021) (Figure 2.5). Much of the T microbiome is made distinct from the RW and LT by the prevalence of families at <5% individual abundance and the unclassified families in the Oxyphotobacteria class, Ignavibacteria class, Gammaproteobacteria class, and "Other" group. The T being the last stage of water reclamation may explain these distinctions in the T microbiome. Water undergoes filtration and UV treatments from the RW, then more UV treatments, biological filters, and mechanical filters in an intermediate tower stage before reaching the mechanical filters and UV treatment on the T. Water cleaning in RASs aims to maintain consistent water quality, such as UV sterilization reducing microbial density (Laroche *et al.*, 2018, Attramadal *et al.*, 2021, Dahle *et al.*, 2022). Though the T microbiome is likely sourced by the RW microbiome, the water reclamation process may explain the lesser abundances of taxa observed in the T microbiome compared to the RW and LT.

As the only compartment not recycled in the RAS loop and consistently replenished by new algae, the AH microbiome is the most distinctive of the four compartments. Though Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia are, like the within-loop compartments, the top three most abundant classes, they are present in lesser abundance compared to the RW, T, and LT microbiomes (Figure 2.5). Additionally, the class Oxyphotobacteria consisted of 25.6% to 30.2% of the AH microbiome, and an unclassified member of the class had the highest individual relative abundance at the family level. Fewer families in general appear in the AH microbiome compared to the within-loop compartments, and those in higher abundances (Oxyphotobacteria, Alteromonadaceae, Flavobacteriaceae, and Thalassospiraceae) are also present in higher abundances in the LT microbiome than in the RW or T. Since the

AH supplies the reared larvae's diet and is added directly into the LTs, it is evident the AH microbiome was another source for the LT microbiome.

Differences between compartments are explained by differences in compartment design and/or function, namely the presence of larvae in the LTs, the new water and filtration in the RW, the T being the final stage of water cleaning, and AH existing outside the RAS loop. Based on these distinctions in compartment design/function and observed distinctions in microbial composition, each compartment applied selective pressures on the RAS microbiome. It is likely the composition of the RW microbiome is dependent on the composition of the LT microbiome and source water, the T microbiome is dependent on the composition of RW microbiome, and the LT microbiome is supplied by the T microbiome, the larvae being reared, and the AH microbiome feeding the larvae. A previous study on RAS microbiomes noted each RAS compartment contained a unique "microenvironment" (Bartelme *et al.*, 2019). Though this study also notes differentiations between compartments, the influence of one compartment on the other cannot be denied.

Though prior studies of RAS microbiomes have noted a connection between stability and diversity (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Asmani *et al.*, 2016), the present study found no consistent connection between high diversity and increased stability. The T had the highest average stability in terms of successive UniFrac distances with the least deviation over time (0.30 ± 0.08 overall) (Figure 2.18) compared to average stability in the RW (0.37 ± 0.10 overall) (Figure 2.15), LT (0.33 ± 0.14 overall) (Figure 2.21), and AH (0.37 ± 0.19 overall) (Figure 2.24). The T also had the highest average diversity (6.00 ± 0.52 overall) compared to the RW (4.99 ± 0.34 overall), LT (4.14 ± 0.63 overall), and AH (2.80 ± 0.11 overall) (Table 2.5). However, the LT had

higher average stability and lower average diversity compared to the RW. Over time, trends in stability and diversity per compartment rarely aligned. For example in the RW (Figures 2.4,2.15), both stability and diversity were initially high in 2021, but the highest moment of instability in the RW observed in 2021 did not coincide with a dip in diversity. Contextualizing stability in terms of the starting sample in 2019 (Figures 2.16, 2.19, 2.22) confirms there is no connection between stability and diversity in that trends in stability across the within-loop compartments do not align with trends in diversity. Attramadal *et al.* (2014) suggested the high diversity observed in their RASs was indicative of a mature microbial community and, with that, a more stable microbiome compared to FT systems. While this may be the case when comparing different aquaculture systems, the present study did not observe high diversity as an identifier of a more stable microbiome or vice versa.

Stability was quantified by the MWA in terms of distances between successive samples and distances from the starting sample in 2019 to assess the impacts of time on RAS microbial stability. In terms of the former, stability only significantly increased from 2019 to 2021 in the T ($p< 0.01$, Figure 2.18) and from 2020 to 2021 in the RW ($p=0.01$, Figure 2.15). In terms of the latter, only a decrease in stability between 2019 and 2021 in the T was significant ($p=0.04$, Figure 2.19). Neither quantification method confirms the hypothesized increase in RAS stability over a three-year period. There was no trend of increased instances of stability or instability over time in the RAS. However, some peaks and drops in stability aligned by day of sampling across compartments, suggesting another factor other than time influences microbial stability in each compartment simultaneously.

Additionally, significant changes in stability over time were detected by the STARS algorithm primarily at the end of 2021 (Figures 2.17, 2.20, 2.23, 2.26), implying the end of the overall experiment was significantly more stable (in the RW, T, and AH) or more unstable (in the LT) compared to the average trends in stability over the course of three runs. Few shifts in stability were detected outside of the final four to nine time points, and though those that were (one or two in within-loop compartments) did prelude rises or falls in stability, the detection algorithm did not identify observed shifts in community composition, such as the introduction of the BS tank in the RW in 2019. The RAS microbiome has been shown to stabilize across multiple spawning periods (Dahle *et al.*, 2022), but the results of this study indicate a more complicated relationship between microbial stability and time. Though stability was not achieved by 2021, there may be other factors influencing fluctuations in stability that were more apparent in 2021, such as larval survival (explored in chapter 3).

2.4.2 Brood Introductions and Removals Have a Minimal Effect on the RAS Microbiome

There is a relationship between oyster larval and water microbiomes that is evident in literature (Asmani *et al.*, 2016, Laroche *et al.*, 2018, Arfken *et al.*, 2021). This relationship, however, has little to no impact on the LT microbiome nor the microbiomes of the RW and T. Whether larvae were present or absent in the system had no statistically significant impact on the RAS microbiomes (2019 being the exception is discussed later) (Table 2.16). While not statically significant, there is an observed connection between periods of instability and exchanges of larval broods. The longer a brood remained in the system, the more stable the microbiomes of the within-loop compartments became.

Some, though not all, exchanges of one brood for another brood or for a period of larval absence are marked by a spike in instability. For example, the transition between 2021's brood 6 for brood 7 (days 73 to 80) matches the highest moment of instability in the RW and LT as well as a rise in instability in the T (Figures 2.15,2.18,2.21). A similar rise and fall in instability occur in the transitions between all four broods in 2020 and in the introduction of broods 3, 5, 8, 9, 10, and 11 in 2021. RAS microbial stability appeared to fluctuate in a pattern set by larval production. The pattern is not exact, however, which suggests another feature in production may be influencing microbial stability. Larval survival may be this feature, and the connection between larval health and RAS microbial stability will be investigated in chapter 3. What can be said is instances of instability in the RW and T microbiomes align with larval rearing even though there is no statistical difference in the presence or absence of larvae.

2.4.3 Changes in System Design Shift Community Composition

Confounding factors arising in 2019 caused a mortality event in the RAS and shifts in the microbiomes of the RW and T. These factors included an additional tank added during sampling, additional microbial source in the form of adult oysters, trace metal contamination, and trace metal treatments. Changes to aquaculture system designs, such as the degree of water recirculation and additional compartments, have been noted to impact microbial community composition (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Asmani *et al.*, 2016, Almeida *et al.*, 2021). This study found the addition of a new compartment and all its conflated factors also impacted community composition, specifically in the RW and T. In the RW, a signature community emerged during the period when the BS was present in the system (Figures 2.29,2.30). Community structure

was found to be distinct from any other period in 2019 as well as 2020 and 2021 (Table 2.18). The shifted community persisted shortly after the BS was removed before eventually resembling the community composition in 2020. Treatment administration then caused another shift in the community that was not permitted the opportunity to potentially righten. In the T, a similar effect occurred though not to the same degree (Figures 2.31,2.32, Table 2.19). A slight shift in community structure occurred when the BS was present, but the microbiome returned to the pre-BS structure once the BS was removed. Treatments administered to the system had a similar effect on community composition as seen in the RW. The LT microbiome, however, displayed no distinct shifts in response to events in 2019 (Figures 2.33,2.34, Table 2.20).

The compartments' distinct reactions to the BS event shed light on the ways the RAS microbiomes respond to disturbances using different mechanisms of stability, namely resilience and resistance (Allison & Martiny, 2008). The RW and T microbiome displayed resilience in that both compartments experienced shifts in microbial composition that were, once the disturbance ended, mostly returned to their original form. The LT microbiome exhibited resilience to changes, maintaining community structure despite the disturbances. These mechanisms are further evident in trends in stability over time (Figures 2.15,2.18,2.21). The RW microbiome became unstable shortly after the BS was added to the system and stabilized after the BS was removed. Again, shortly after system treatments began, the RW microbiome experienced instability before gradually stabilizing through the rest of sampling in 2019. Similarly, the T microbiome experienced its highest moment of instability when the BS was introduced, and stability declined the longer the BS was present. After the BS was removed, the T microbiome gradually

stabilized. The LT microbiome also experienced a decline in stability after the BS was introduced but generally stabilized over time regardless of the rest of 2019's events.

Through the events in 2019, the RAS microbiome proved how adept it was at handling a massive disturbance that caused production losses. Simply returning the RAS to standard operations allowed the RW and T microbial communities to return to a composition similar to their pre-disturbance structure while the LT resisted the disturbance altogether. These insights into the dynamics of the RAS microbiome, however, were not known at the time of the event in 2019. The decision to restart the system after sampling ended was a response to the trace metals detected and the larval mortality the system experienced.

Overall differences between the RAS microbiomes in 2019 and the microbiomes in 2020 and 2021 reflect the system restart after 2019 and echo a similar observation made by Dahle *et al.* (2022). Dahle and company observed the microbiome of an RAS across three batches of salmon fry in which the system was drained and the biofilters disinfected after the first batch only. It was then observed the RAS microbiome during the first batch of salmon fry was different from the microbiome during batches two and three. They posited these results were due in part to the system restart. The present study agrees with the finding, having observed that the RW and T microbiomes distinguished the communities 2019 from those of 2020 and 2021. Draining and cleaning a closed system between spawning periods not only restarts the physical system but also the microbial community established during the prior period of operation.

2.4.4 Limitations

Certain components of the present study created limitations in the sample collection method, sequencing effort, and downstream analyses. These limitations do not diminish the knowledge gained but do offer points of improvement that should be acknowledged. It is recommended the discussed improvements be implemented in future studies of RAS microbiome dynamics.

The RAS studied aimed to optimize the production of *C. virginica* larvae, and though investigating the microbial community of this closed system was a primary focus, decisions were made in the interest of larval production. For example, the system reset post-2019 sampling was a decision made with salvaging production efforts in mind. Cleaning out the possible contaminants introduced by the BS tank by draining and replenishing the system reset the microbiome established over the run. Further explorations into how the RAS microbiome responses to and/or recovers from disturbance events may offer an optimal course of action for handling disturbances during production other than resetting the system.

Additionally, though the RAS was not broken down between production runs (2019 being an exception), sampling of the water was only done during larval production, and no sampling was done on the source water. Brood rearing in the system began before sampling and lasted after sampling ended for all three runs. The effects of larval presence in the system were inferred using periods of larval absence, but the dynamics of the RAS microbiome in the absence of larvae, namely during the off-season, are unknown. This gap in the time series created a limitation in establishing a baseline community in this closed system and prevented further exploration into the impacts larvae have on the

closed system's microbiome. A lack of replicates and inconsistent sampling days across runs also created gaps in the time series. Having no replicate samples forced individual samples to represent a unique time and place with no means of recovering that time and place should the sample be lost, as was seen in the removal of outliers (Tables 2.1, 2.2). Inconsistency in sampling days between runs coupled with sampling beginning during an arbitrary point in spawning meant direct comparisons between days of each run were not possible. The lack of source water samples also limited the study's ability to make comparisons to a source community. Understanding the community that establishes and replenishes the closed system may provide a pseudo-control variable from which all RAS compartments may be compared, as has been done in prior RAS microbiome studies (Attramadal *et al.*, 2014, Bartelme *et al.*, 2019). It is recommended for future studies of RAS microbiomes to establish a stricter sampling effort that includes periods when no larval rearing is occurring and the source water.

The sequencing effort was shown to have underrepresented community composition in all four RAS compartments (Table 2.5). The chosen primers for this study targeted the V6-V8 regions of the 16S rRNA gene due to the regions' high resolution of taxa (Parada *et al.*, 2016) and higher distinctions between lower taxonomic levels (Bukin *et al.*, 2019). However, the chosen regions did not adequately capture the microbial communities in an RAS, especially the microbiomes of the within-loop compartments. It is recommended for future studies to use a different sequencing approach, such as using primers targeting the V4-V5 regions previously used in studies of RASs (Attramadal *et al.*, 2014, Bakke *et al.*, 2017, Bartelme *et al.*, 2019, Dahle *et al.*, 2022) and oysters (Pierce *et al.*, 2016, Stevick *et al.*, 2019, Arfken *et al.*, 2021, Pimentel *et al.*, 2021).

The effects of both the sampling and sequencing methods impacted downstream analyses. The sequencing method inadequately capturing community composition limits the reliability of lower-level taxonomic assessments, such as visualizing community composition (Figure 2.5) and later analyses of phylotypes in chapter 3. The sampling method lacking a means of establishing a baseline community limits the reliability of stability analyses, namely the detection of regime shifts (Figures 2.17, 2.20, 2.23, 2.26). Smoothing of stability data allowed for a higher resolution of regime shifts. However, the STARS algorithm more reliably detects shifts when a threshold or baseline is established in the time series. The calculations use a regime average to establish a threshold for acceptable values, or those that do not significantly deviate from the regime. Values falling outside the acceptable deviation is a candidate for a regime shift, and the RSI for that value is calculated based on a specified regime length. Significant deviations from the average constitute a shift. It is evident, then, that shifts are better detected when a sample's deviation from the norm is clearer. Though smoothing aided this endeavor, establishing a clearer baseline would better highlight regime shifts in the RAS microbiome during production. This limitation may explain why shifts were more likely to be detected at the end of the experiment for all compartments. This study marks the first time this method has been used on microbial stability over time for the purpose of detecting significant changes in stability in RASs. Though this research provides an initial framework for future investigations of closed system microbiomes, it does so with the understanding more research must be done on time series involving microbial stability to determine the method's reliability.

2.5 Conclusions

The RAS microbiome was influenced by both space and time. Individual compartments shared taxa with each other and the diet the larvae were fed (AH). However, differences in compartment function, such as water cleaning or larval rearing, resulted in distinct distributions of taxa across all four compartments. Over time, physicochemical conditions generally become steadier in both the RW and T, and the impact physicochemical conditions had on community composition lessened with the decline in variability. No connection between stability and diversity was observed, nor was a trend of increased stability over time observed in the RAS microbiome. Though the presence or absence of larvae in the system held no statical significance, some instances of instability in the RW and T microbiomes aligned with larval brood introductions. Additionally, the introduction of adult oysters and change in system design in 2019 shed light on the RAS microbiome's ability to withstand disturbance events.

CHAPTER III – DETERMINE CONNECTIONS BETWEEN OYSTER LARVAE SURVIVAL AND RAS MICROBIOME STABILITY

3.1 Connections Between Larval Survival and the RAS Microbiome

That the oyster larval microbiome and their health are influenced by their water microbiome has been established (Lokmer & Wegner, 2015, Lokmer *et al.*, 2016, Lokmer *et al.*, 2016, Green *et al.*, 2019). Chapter 2 determined a connection between larval introductions and RAS microbial stability. How the RAS microbiome is influenced by the health of oyster larvae is not defined. The dynamics of a closed system may exhibit patterns of change associated with larval mortality, and features of the RAS microbiome may be associated with brood outcomes. It is hypothesized the stability of the RAS microbiome will connect with larval survival rates, indicating a relationship between the microbiome and production outcomes. Should the null of no difference be accepted, there is no connection and, thus, no relationship between the microbiome and production outcomes.

3.2 Methods and Materials

3.2.1 Larval Survival Data

Larval broods were spawned in-house (TCMAC) with oysters bred from selected broodstock. The resulting eggs were counted, and hatched larvae were counted two days later. The ratio of amount hatched to the number of eggs was used to calculate the percentage hatched. After day 2, larvae were retrieved during LT drain downs to be counted every two days until the larvae were removed from the system, either when larvae were large enough to be harvested or the brood was no longer viable. To note, LT drain downs involved moving LT water directly to the RW and minimizing water loss.

The number of larvae (counts) from each tank was estimated by adding the larvae into a bucket or beaker with clean culture water to a volume that condensed the larval density to a point at which a 100 μL aliquot contained 50 to 100 larvae. Three 100 μL aliquots were collected from the counting container and added to a Sedgewick Rafter cell for counting using a compound microscope. The total number of larvae in the tank was extrapolated volumetrically from the average number of larvae in the three 100 μL aliquots. Percent survival was calculated from the ratio of the amount harvested to the amount hatched on Day 2. Nonviable broods were labeled "Terminated" while surviving broods were "Harvested". Some broods were reared at the same time regardless of each brood's performance. So, water samples taken when broods of different outcomes were reared were labeled "Overlapping". If an LT sample and a reared brood were documented to be in the same location (one of four larval tanks), then the LT sample was assigned the brood outcome of its corresponding brood.

3.2.2 Statistical Techniques

To determine connections between brood outcomes and the RAS microbiome, data was statistically analyzed using the PRIMER7 software. The weighted UniFrac matrix imported into PRIMER7 was used to visualize the similarity and differences between samples with non-metric multi-dimensional scaling (nMDS). This was used to visually orient samples as two-dimensional plots to aid in understanding how similar or different the microbial community composition is between samples associated with harvested broods, terminated broods, and overlapping outcomes. Additionally, samples were categorized by survival rates (5% intervals) and terminations to visualize connections between percent survival and microbial composition.

A BEST procedure, as defined in chapter 2, was used to assess the correlation between the RAS microbiome and larval performance in terms of brood outcome and survival rate. Survival rates were already numerical values, but brood outcomes as a categorical factor required coding. Samples associated with Harvested broods, Terminated broods, and larval absence were coded as 2, 1, and 0, respectively.

A SIMPER analysis was then used to analyze the contribution of phylotypes to brood outcomes if the BEST procedure indicated a correlation between a compartment's microbiome and larval performance. The goal of the SIMPER analysis was to determine which specific phylotypes contributed to changes in microbial community structure based on brood outcomes. The analysis calculated the average (dis)similarity of samples within/between groups by generating a Bray-Curtis similarity matrix from relative abundance data per compartment. Average abundance, (dis)similarity, ratio of (dis)similarity to standard deviation of contribution, and percent contribution of each phylotype were then calculated. These calculations were defined based on user selected groups. Hence, the SIMPER routine cannot be used as a significance test, just as a means of investigating drivers of differences in identified clusters via phylotype contribution. The results of the analysis described specific phylotypes within and between defined groups, here brood outcomes. The relative abundance of selected high contributing phylotypes were then visualized on nMDS plots, where symbols were scaled to relative abundance and colored by survival rate intervals, to visualize the influence of phylotypes on larval survival and the orientation of samples.

3.3 Results

3.3.1 Brood Outcomes Across Three Runs

As summarized in Table 3.1, a total of 35 broods were reared in the RAS during the experiment's timeframe. Of those, 19 broods were harvested while 16 were terminated. Twenty broods were spawned in 2019 with only 8 succeeding in a harvest. Four broods were spawned in 2020 with only one terminated. Eleven broods spawned in 2021 with only three terminated. Overall, 2020 and 2021 had the highest success rates in terms of successive to unsuccessful broods. Year 2021 had the highest average survival rate of about 23.9% from successful broods compared to 2020's 8.4% and 2019's 10.4%.

Table 3.1 *Summary of broods reared in each run*

Data includes broods eared successfully (harvested = H) and unsuccessfully (terminated = T) in each run. Percent survival of harvested broods calculated from the hatched count and final larval count at time of harvest. Production runs vary in brood numbers due to the variation of oyster larvae production outcomes in the RAS over three runs.

2019 Broods	Outcome	% Survival	2020 Broods	Outcome	% Survival	2021 Broods	Outcome	% Survival
1	H	28.57	1	T	NA	1	H	24.81
2	T	NA	2	H	2.74	2	H	24.26
3	H	10.50	3	H	18.70	3	H	0.15
4	H	16.01	4	H	3.77	4	T	NA
5	T	NA				5	H	22.51
6	T	NA				6	T	NA
7	T	NA				7	H	43.80
8	T	NA				8	H	18.08
9	T	NA				9	T	NA
10	T	NA				10	H	20.96
11	H	0.67				11	H	36.28
12	H	1.72						
13	T	NA						
14	H	4.81						
15	T	NA						
16	H	10.39						
17	T	NA						
18	T	NA						
19	H	10.39						
20	T	NA						

3.3.2 RW Analysis

The RW dataset was visualized via an nMDS plot in Figure 3.1 to determine orientation of samples to each other based on brood outcome (left) and survival rates (right). Samples taken during harvested broods in 2019 are primarily grouped together or with other harvested samples from other runs. Most 2019 terminated samples are tightly knit in the bottom center of the plot. Samples with overlapping outcomes in 2019 are oriented mostly with harvested 2019 samples. Samples taken when larvae were absent were mostly associated with harvested and overlapping samples save one outlier sample. Harvested samples from 2020 mostly orient above samples from 2019. The single terminated sample from 2020 is included in an outlier cluster at the bottom left of the plot with two other 2020 samples and a 2021 terminated sample. No broods overlapped in 2020. 2020 samples taken when larvae were absent mostly group with harvested samples from 2020 and 2021. Harvested samples from 2021 primarily occupy the top of the plot. One outlier cluster at the top left of the graph is comprised mostly of harvested 2021 samples. The single terminated 2021 sample is included in the bottom left outlier cluster. Overlapping 2021 samples group entirely with harvested 2021 samples.

Further visualizing RW samples by larval survival rates, samples taken during low performing broods (survival rates <10%) largely orient together and with terminated samples. Samples associated with high performing broods (>30%) generally orient in the top half of the plot. Samples associated with survival rates between 10% and 30% are mostly spread across the plot near samples of high and low performing broods.

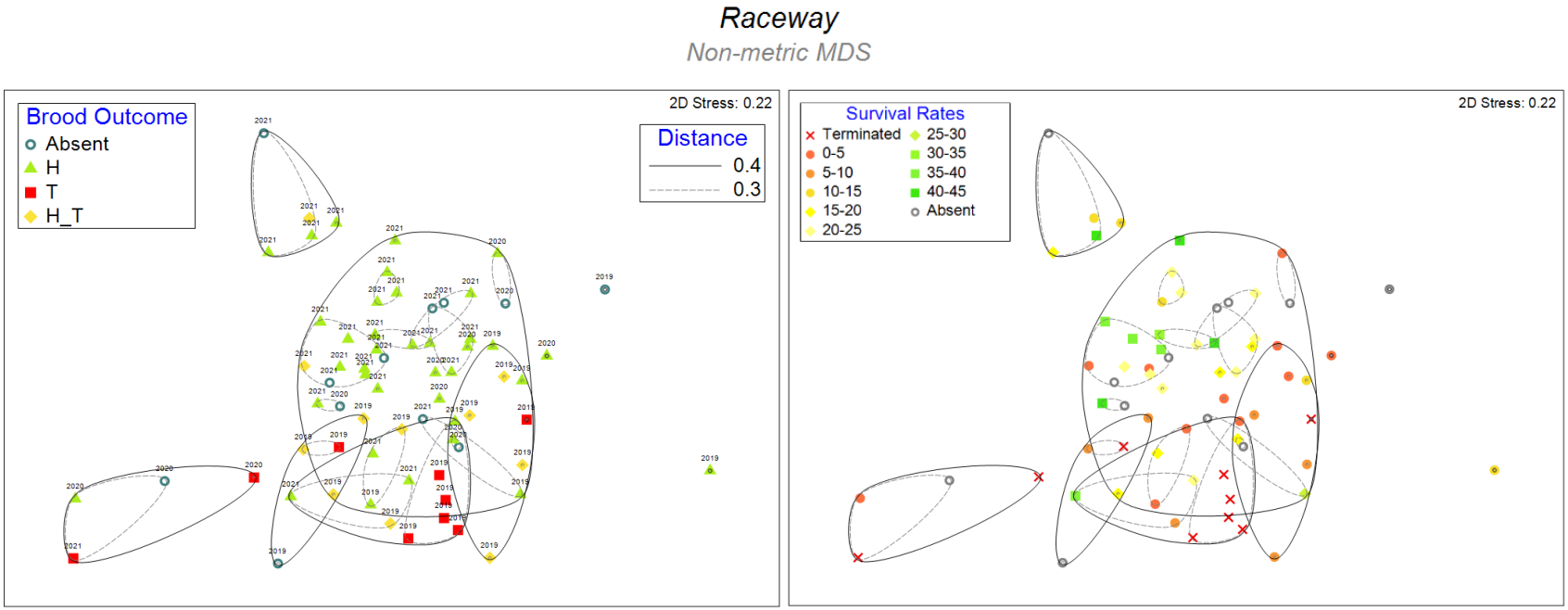

Figure 3.1 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the RW during all three runs visualizing larval performance*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by brood outcomes (left) and percent survival (right) for visualization. Outcomes consist of harvested (H), terminated, overlap of harvested and terminated broods (H_T), and absence of larvae. Numbers above each symbol indicate the run the sample was taken. Categories of survival define Terminated, 0-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-30%, 30-35%, 35-40%, 40-45%, and the absence of larvae. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).

A BEST procedure was performed on larval survival rates and brood outcomes in the RW in each run to determine the degree to which larval performance shapes the microbiome (Table 3.2). In 2019, survival rates had a higher individual correlation with community composition compared to brood outcomes (0.114 and 0.015, respectively). A combination of variables cumulated in a correlation of 0.113. In 2020, brood outcomes had a higher individual correlation with community composition compared to survival rates (0.123 and 0.006, respectively). A combination of variables cumulated in a correlation of 0.006. In 2021, brood outcomes had a higher individual correlation with community composition compared to survival rates (0.256 and 0.110, respectively). A combination of variables cumulated in a correlation of 0.111. Compared to water quality parameters in the RW (see chapter 2.3.5), larval performance in terms of survival rates and brood outcomes did not have as strong an influence on the RW microbiome in 2019 and 2020. In 2021, however, correlations between larval performance and the RW microbiome were higher than that of water quality parameters. Comparing results across all three runs, survival rates had the highest correlation in 2019 and the lowest in 2020 while brood outcome had the highest correlation in 2021 and lowest in 2019. These results suggest a connection between larval performance and the RW microbiome exists and strengthened over time.

Table 3.2 *BEST procedure matching the RW's weighted UniFrac distance matrix in correlation with larval survival rates and brood outcomes*

BEST Results: Larval Performance in the Raceway								
Correlation method: Spearman rank								
Maximum number of variables: 2								
Analyze between: Samples from RW Weighted UniFrac Distance Matrix								
Resemblance measure: D1 Euclidean distance								
Run 2019			*Run 2020*			*Run 2021*		
Number of variables: 1			*Number of variables: 1*			*Number of variables: 1*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.114	% Survival	1	0.123	Brood Outcome	1	0.256	Brood Outcome
1	0.015	Brood Outcome	1	0.006	% Survival	1	0.110	% Survival
Number of variables: 2			*Number of variables: 2*			*Number of variables: 2*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
2	0.113	% Survival, Brood Outcome	2	0.006	% Survival, Brood Outcome	2	0.111	% Survival, Brood Outcome
Best result for each number of variables			*Best result for each number of variables*			*Best result for each number of variables*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.114	% Survival	1	0.123	Brood Outcome	1	0.256	Brood Outcome
2	0.113	% Survival, Brood Outcome	2	0.006	% Survival, Brood Outcome	2	0.111	% Survival, Brood Outcome
Best results			*Best results*			*Best results*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.114	% Survival	1	0.123	Brood Outcome	1	0.256	Brood Outcome
2	0.113	% Survival, Brood Outcome	1	0.006	% Survival	2	0.111	% Survival, Brood Outcome
1	0.015	Brood Outcome	2	0.006	% Survival, Brood Outcome	1	0.110	% Survival

To investigate specific phylotypes shaping the RW microbiome in relation to brood outcome, a SIMPER analysis was performed using relative abundance data on the samples included in Figure 27. In 2019 (Table 3.3), samples taken when larvae were absent were predominantly shaped by a member of the Rhodobacteraceae family (37% contribution), the Flavobacteriaceae family (12%), and an uncultured phylotype from the Leptotrichiaceae family (9%). Samples taken during harvested broods were predominantly shaped by a member of the Rhodobacteraceae family (16% contribution), chloroplast (9%), and a member of the Alteromonadaceae family (6%). Samples taken during terminated broods were dominated by a member of the Rhodobacteraceae family (58% contribution). Samples taken when both harvested and terminated broods were reared were predominantly shaped by a member of the Rhodobacteraceae family (26% contribution), two members of the Alteromonadaceae family (cumulatively 14%), chloroplast (6%), and a member of the class Gammaproteobacteria (5%). Overall, 2019 samples associated with terminated broods had the highest average similarity (37.54) while absent and harvested samples had the lowest average similarities (17.02 and 17.80, respectively).

The pairwise test (Table 3.4) displays phylotypes contributing to dissimilarities between samples associated with different brood outcomes in 2019. Abundances of the member of the family Rhodobacteraceae contributed the most to differences between absent and terminated samples (17%), overlapping and terminated samples (18%), and harvested and terminated samples (19%). Abundances of members of the families Saccharospirillaceae, Flavobacteriaceae, Thalassospiraceae, Alteromonadaceae, as well as an uncluttered Owenweeksia species and chloroplast contributed <11% to dissimilarity

between the three pairs of interest. Overall, the pairwise test indicated absent and terminated samples had the highest average dissimilarity (83.60) and absent and overlapping samples had the lowest average dissimilarity (74.94).

In 2020 (Table 3.5), samples taken when larvae were absent were predominantly shaped by members of the Rhodobacteraceae family (10% contribution), the Alteromonadaceae family (5%), the Thalassospiraceae family (5%), the Pseudoalteromonadaceae family (3%), as well as two members of the class Oxyphotobacteria (cumulatively 6%). Samples taken during harvested broods were predominantly shaped by members of the Flavobacteriaceae family (6% contribution), the Rhodobacteraceae family (6%), the Alcanivoracaceae family (5%), as well as an undefined member of the phylum Planctomycetes (4%), two members of the class Gammaproteobacteria (cumulatively 7%), and chloroplast (3%). No test was performed for the terminated group as only one sample belongs to the group. No broods overlapped during rearing in 2020. Overall in 2020, harvested samples were more similar (average 30.30 similarity) than absent samples (average 18.26 similarity). Given no test was possible for the terminated group, no pairwise test was performed.

In 2021 (Table 3.6), samples taken when larvae were absent were predominantly shaped by members of the families Cellvibrionaceae (23% contribution), Rhodobacteraceae (7%), Spongiibacteraceae (6%), Moraxellaceae (4%), Flavobacteriaceae (4%), and Rhizobiaceae (3%) as well as a member of the order Chitinophagales (3%). Samples taken during harvested broods were predominantly shaped by members of the families Cellvibrionaceae (12%), Rhodobacteraceae (8%), Burkholderiaceae (5%), Alteromonadaceae (4%), Flavobacteriaceae (4%), and

Spongiibacteraceae (3%) as well as a member of the order Chitinophagales (6%). No test was performed for the terminated group as only one sample belongs to the group. Samples taken during overlapping brood outcomes were predominantly shaped by members of the families Spongiibacteraceae (19%), Cellvibrionaceae (13%), Moraxellaceae (6%), Vibrionaceae (4%), Flavobacteriaceae (4%), and Nocardiaceae (3%) as well as an uncultured member of the phylum Planctomycetes (3%). Overall, harvested samples had the highest average similarity (30.98) and overlapping samples had the lowest average similarity (18.18).

The pairwise test (Table 3.7) displays phylotypes contributing to dissimilarities between samples associated with different brood outcomes in 2021. Abundances of the families Alteromonadaceae (41%), Cellvibrionaceae (5%), and Thalassospiraceae (4%) as well as a member of the order Chitinophagales (4%) contributed to differences between harvested and terminated samples.

The member from family Rhodobacteraceae was a reoccurring contributor to similarities and dissimilarities between samples associated with varying brood outcomes in all three runs. Members of the families Alteromonadaceae, Thalassospiraceae, Flavobacteriaceae, Saccharospirillaceae, and Cellvibrionaceae contributed to differences in the community of harvested and terminated samples in 2019 and 2021.

Table 3.3 *Similarity Percentage (SIMPER) assessment of the contributions of phylotypes on brood outcomes in the RW in 2019*

SIMPER: Brood Outcomes in the Raceway in 2019					
Phylotypes	**Av.Abund**	**Av.Sim**	**Sim/SD**	**Contrib%**	**Cum.%**
Group Absent					
Average similarity: 17.02					
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.09	6.33	SD=0!	37.16	37.16
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	2.02	SD=0!	11.86	49.02
Fusobacteria, Fusobacteriia, Fusobacteriales, Leptotrichiaceae, uncultured	0.02	1.52	SD=0!	8.9	57.92
Group H_T					
Average similarity: 28.66					
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.13	7.5	1.46	26.17	26.17
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.04	2.08	1.04	7.25	33.42
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.04	1.8	0.95	6.27	39.69
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.05	1.73	0.66	6.03	45.72
Proteobacteria, Gammaproteobacteria	0.02	1.34	1.03	4.68	50.4
Group H					
Average similarity: 17.80					
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.06	2.83	1.08	15.9	15.9
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.03	1.6	1.08	9.01	24.91

Table 3.3 (continued).

Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.03	1.16	0.79	6.49	31.4
Bacteroidetes, Bacteroidia, Flavobacteriales, Cryomorphaceae, uncultured, uncultured Owenweeksia sp.	0.05	0.84	0.69	4.73	36.13
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas, Ambiguous taxa	0.01	0.81	0.99	4.56	40.69
Bacteroidetes, Bacteroidia, Chitinophagales, Saprospiraceae, Phaeodactylibacter, uncultured bacterium	0.02	0.7	0.76	3.95	44.65
Proteobacteria, Deltaproteobacteria, Myxococcales, Haliangiaceae, Haliangium, uncultured bacterium	0.01	0.68	1.3	3.84	48.48
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.62	0.74	3.5	51.99
Group T					
Average similarity: 37.54					
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.35	21.67	1.1	57.74	57.74

Table 3.4 *Similarity Percentage (SIMPER) assessment on observed differences in the relative abundances of phylotypes between brood outcomes in the RW in 2019*

SIMPER Pairwise Test: Brood Outcomes in the Raceway in 2019						
Phylotypes	**Av.Abund**	**Av.Abund**	**Av.Diss**	**Diss/SD**	**Contrib %**	**Cum.%**
Groups Absent & H_T						
Average dissimilarity = 74.94	*Group Absent*	*Group H_T*				
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae, uncultured	0.18	0	9.09	0.97	12.13	12.13
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.09	0.13	3.89	1	5.19	17.32
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.09	0.05	3.85	1.31	5.14	22.46

Table 3.4 (continued).

Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.06	0.04	2.68	1.37	3.58	26.04
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.01	0.05	2.48	0.45	3.31	29.35
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Sphingobium	0.02	0.04	1.79	0.98	2.38	31.74
Bacteroidetes, Bacteroidia, Cytophagales, Cytophagaceae, Cytophaga, uncultured bacterium	0	0.03	1.52	0.62	2.03	33.77
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.02	0.04	1.52	1.13	2.02	35.79
Cyanobacteria, Oxyphotobacteria, Chloroplast, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.01	0.03	1.49	1	1.98	37.78
Bacteroidetes, Bacteroidia, Chitinophagales, Chitinophagaceae, Lacibacter	0.03	0	1.48	0.97	1.98	39.75
Proteobacteria, Alphaproteobacteria, Caulobacterales, Caulobacteraceae, Caulobacter, Ambiguous taxa	0.03	0.01	1.4	1.01	1.87	41.62
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae, Ruegeria	0	0.03	1.25	0.85	1.67	43.29
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Flavobacterium	0	0.02	1.15	0.61	1.54	44.83
Proteobacteria, Deltaproteobacteria	0.03	0.01	1.09	1.52	1.45	46.28
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.02	1.06	1.57	1.42	47.7
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas, Ambiguous taxa	0	0.02	1.01	0.5	1.35	49.05
Bacteroidetes, Bacteroidia, Flavobacteriales, Cryomorphaceae, uncultured, uncultured Owenweeksia sp.	0.01	0.01	0.96	0.94	1.28	50.32

Table 3.4 (continued).

Groups Absent & H						
Average dissimilarity = 81.99	*Group Absent*	*Group H*				
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae, uncultured	0.18	0	9.09	0.96	11.09	11.09
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.09	0.03	3.79	1.27	4.62	15.71
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira, Ambiguous taxa	0.02	0.06	3.46	0.57	4.22	19.93
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.09	0.06	2.79	1.57	3.41	23.33
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.06	0.03	2.79	1.3	3.4	26.73
Bacteroidetes, Bacteroidia, Flavobacteriales, Cryomorphaceae, uncultured, uncultured Owenweeksia sp.	0.01	0.05	2.36	0.67	2.87	29.61
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.01	0.03	1.86	0.51	2.26	31.87
Proteobacteria, Deltaproteobacteria, Bdellovibrionales, Bacteriovoracaceae, Peredibacter, uncultured bacterium	0.01	0.03	1.58	0.53	1.93	33.8
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.03	1.57	1.45	1.91	35.71
Bacteroidetes, Bacteroidia, Chitinophagales, Chitinophagaceae, Lacibacter	0.03	0	1.48	0.96	1.81	37.52
Proteobacteria, Alphaproteobacteria, Caulobacterales, Caulobacteraceae, Caulobacter, Ambiguous taxa	0.03	0	1.4	1.08	1.71	39.23
Proteobacteria, Deltaproteobacteria	0.03	0	1.25	1.76	1.53	40.76
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas, Ambiguous taxa	0	0.02	1.22	0.47	1.48	42.24
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Aequorivita	0	0.02	1.21	0.43	1.47	43.71
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Sphingobium	0.02	0.02	1.19	1.02	1.45	45.16

Table 3.4 (continued).

Bacteroidetes, Bacteroidia, Sphingobacteriales, Sphingobacteriaceae, Pedobacter	0	0.02	1.08	0.47	1.32	46.48
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.02	0.01	0.97	1.43	1.18	47.67
Bacteroidetes, Bacteroidia, Chitinophagales, Saprospiraceae, Phaeodactylibacter, uncultured bacterium	0	0.02	0.96	0.98	1.17	48.84
Fusobacteria, Fusobacteriia, Fusobacteriales, Leptotrichiaceae, uncultured	0.02	0	0.93	4.14	1.14	49.97
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, NS3a marine group	0	0.02	0.91	0.43	1.1	51.08
Groups Absent & T						
Average dissimilarity = 83.60	*Group Absent*	*Group T*				
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.09	0.35	14.33	1.63	17.14	17.14
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae, uncultured	0.18	0	9.09	0.96	10.87	28.02
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.09	0.01	4.11	1.12	4.92	32.93
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, NS3a marine group	0	0.06	2.95	1.08	3.53	36.47
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.06	0.03	2.85	1.25	3.41	39.88
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.01	0.03	1.64	0.49	1.96	41.84
Bacteroidetes, Bacteroidia, Chitinophagales, Chitinophagaceae, Lacibacter	0.03	0	1.48	0.96	1.77	43.61
Proteobacteria, Gammaproteobacteria, Alteromonadales, Pseudoalteromonadaceae, Pseudoalteromonas	0	0.03	1.4	0.55	1.68	45.29
Proteobacteria, Alphaproteobacteria, Caulobacterales, Caulobacteraceae, Caulobacter, Ambiguous taxa	0.03	0	1.4	0.96	1.68	46.97
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.01	1.34	2.06	1.6	48.57
Proteobacteria, Deltaproteobacteria	0.03	0	1.24	1.7	1.48	50.05

Table 3.4 (continued).

Groups H_T & H						
Average dissimilarity = 77.12	*Group H_T*	*Group H*				
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.13	0.06	4.95	1.15	6.42	6.42
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.05	0.03	3.48	0.6	4.51	10.93
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira, Ambiguous taxa	0	0.06	2.93	0.46	3.8	14.73
Bacteroidetes, Bacteroidia, Flavobacteriales, Cryomorphaceae, uncultured, uncultured Owenweeksia sp.	0.01	0.05	2.48	0.72	3.22	17.95
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.05	0.03	2.09	1.24	2.71	20.66
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Sphingobium	0.04	0.02	1.96	1.02	2.54	23.2
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.04	0.03	1.78	1.27	2.31	25.51
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.04	0.01	1.73	1.04	2.24	27.75
Cyanobacteria, Oxyphotobacteria, Chloroplast, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.03	0.01	1.58	1.06	2.05	29.79
Proteobacteria, Deltaproteobacteria, Bdellovibrionales, Bacteriovoracaceae, Peredibacter, uncultured bacterium	0	0.03	1.54	0.52	2	31.79
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.02	0.03	1.53	1.1	1.99	33.78
Bacteroidetes, Bacteroidia, Cytophagales, Cytophagaceae, Cytophaga, uncultured bacterium	0.03	0	1.52	0.64	1.97	35.75
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae, Ruegeria	0.03	0.01	1.28	1.06	1.66	37.4
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas, Ambiguous taxa	0	0.02	1.27	0.53	1.65	39.06
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Aequorivita	0	0.02	1.21	0.44	1.57	40.62

Table 3.4 (continued).

Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Flavobacterium	0.02	0.01	1.2	0.69	1.56	42.18
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas, Ambiguous taxa	0.02	0.01	1.17	0.69	1.52	43.71
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas	0.02	0.01	1.15	0.92	1.49	45.2
Bacteroidetes, Bacteroidia, Sphingobacteriales, Sphingobacteriaceae, Pedobacter	0	0.02	1.09	0.49	1.41	46.61
Proteobacteria, Gammaproteobacteria, Xanthomonadales, Xanthomonadaceae, Pseudoxanthomonas	0.02	0.01	1	0.71	1.29	47.91
Proteobacteria, Gammaproteobacteria	0.02	0.01	0.97	1.42	1.25	49.16
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, NS3a marine group	0	0.02	0.91	0.44	1.17	50.33
Groups H_T & T						
Average dissimilarity = 76.22	*Group H_T*	*Group T*				
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.13	0.35	13.49	1.7	17.7	17.7
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.05	0.03	3.31	0.57	4.35	22.05
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, NS3a marine group	0	0.06	2.95	1.11	3.87	25.92
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.05	0.01	2.22	1	2.91	28.83
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.04	0.03	2	1.3	2.62	31.46
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.04	0	1.87	1.08	2.45	33.91
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Sphingobium	0.04	0	1.82	0.85	2.39	36.3
Bacteroidetes, Bacteroidia, Cytophagales, Cytophagaceae, Cytophaga, uncultured bacterium	0.03	0	1.52	0.63	1.99	38.29

Table 3.4 (continued).

Cyanobacteria, Oxyphotobacteria, Chloroplast, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.03	0.01	1.5	1.02	1.96	40.25
Proteobacteria, Gammaproteobacteria, Alteromonadales, Pseudoalteromonadaceae, Pseudoalteromonas	0.01	0.03	1.39	0.6	1.83	42.08
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae, Ruegeria	0.03	0.02	1.31	1.19	1.72	43.8
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas, Ambiguous taxa	0.02	0.01	1.21	0.65	1.59	45.39
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas	0.02	0.01	1.14	0.89	1.49	46.89
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Flavobacterium	0.02	0	1.12	0.63	1.48	48.36
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.02	0.01	1.11	1.15	1.46	49.83
Proteobacteria, Gammaproteobacteria	0.02	0.01	1.03	1.39	1.35	51.17
Groups H & T						
Average dissimilarity = 80.79	*Group H*	*Group T*				
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.06	0.35	15.11	1.61	18.7	18.7
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, NS3a marine group	0.02	0.06	3	1.19	3.71	22.41
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira, Ambiguous taxa	0.06	0	2.95	0.46	3.65	26.06
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Saccharospirillaceae, Oleibacter, uncultured bacterium	0.03	0.03	2.59	0.59	3.2	29.26
Bacteroidetes, Bacteroidia, Flavobacteriales, Cryomorphaceae, uncultured, uncultured Owenweeksia sp.	0.05	0.01	2.33	0.65	2.89	32.15
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.03	0.03	1.9	1.18	2.35	34.5
Proteobacteria, Deltaproteobacteria, Bdellovibrionales, Bacteriovoracaceae, Peredibacter, uncultured bacterium	0.03	0	1.47	0.48	1.81	36.31

Table 3.4 (continued).

Proteobacteria, Gammaproteobacteria, Alteromonadales, Pseudoalteromonadaceae, Pseudoalteromonas	0.01	0.03	1.44	0.63	1.78	38.09
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.03	0.01	1.28	1.42	1.58	39.68
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.01	1.26	0.75	1.56	41.24
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas, Ambiguous taxa	0.02	0	1.23	0.5	1.53	42.76
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae, Aequorivita	0.02	0	1.21	0.44	1.49	44.26
Bacteroidetes, Bacteroidia, Sphingobacteriales, Sphingobacteriaceae, Pedobacter	0.02	0	1.08	0.48	1.34	45.6
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Sphingobium	0.02	0	0.95	0.68	1.18	46.78
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae, Ruegeria	0.01	0.02	0.93	1.22	1.15	47.93
Bacteroidetes, Bacteroidia, Chitinophagales, Saprospiraceae, Phaeodactylibacter, uncultured bacterium	0.02	0.01	0.87	1.08	1.07	49
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas	0.01	0.01	0.85	0.68	1.06	50.06

Table 3.5 *Similarity Percentage (SIMPER) assessment of the contributions of phylotypes on brood outcomes in the RW in 2020*

SIMPER: Brood Outcomes in the Raceway in 2020					
Phylotypes	**Av.Abund**	**Av.Sim**	**Sim/SD**	**Contrib%**	**Cum.%**
Group Absent					
Average similarity: 18.26					
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.04	1.88	0.97	10.31	10.31
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.11	0.95	0.95	5.19	15.5
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira	0.05	0.84	0.44	4.58	20.08
Cyanobacteria, Oxyphotobacteria	0.02	0.63	0.92	3.43	23.5
Proteobacteria, Gammaproteobacteria, Alteromonadales, Pseudoalteromonadaceae, Pseudoalteromonas	0.01	0.59	0.89	3.23	26.73
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.02	0.56	0.66	3.05	29.78
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.01	0.5	1.41	2.75	32.53
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Aestuariibacter	0.04	0.48	1.3	2.6	35.14
Proteobacteria, Gammaproteobacteria, Vibrionales, Vibrionaceae, Vibrio	0.01	0.42	2.9	2.32	37.46
Proteobacteria, Gammaproteobacteria	0.01	0.4	1.66	2.18	39.63
Proteobacteria, Gammaproteobacteria, KI89A clade	0.01	0.38	0.52	2.09	41.73
Proteobacteria, Gammaproteobacteria, Xanthomonadales, Xanthomonadaceae, Stenotrophomonas	0.04	0.38	0.75	2.07	43.79
Planctomycetes, OM190	0.01	0.37	0.99	2.05	45.84
Proteobacteria, Alphaproteobacteria, Micavibrionales, uncultured, uncultured bacterium, uncultured bacterium	0.01	0.37	0.53	2.02	47.86
Proteobacteria, Alphaproteobacteria	0.01	0.35	0.76	1.89	49.75
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas, Ambiguous taxa	0.01	0.33	0.9	1.8	51.56

Table 3.5 (continued).

Group H					
Average similarity: 30.30					
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.05	1.72	0.95	5.69	5.69
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.04	1.68	0.96	5.56	11.25
Proteobacteria, Gammaproteobacteria, Oceanospirillales, Alcanivoracaceae, Alcanivorax	0.03	1.4	0.78	4.63	15.88
Planctomycetes, OM190	0.02	1.33	1.82	4.4	20.28
Proteobacteria, Gammaproteobacteria	0.01	1.03	2.08	3.39	23.66
Proteobacteria, Gammaproteobacteria, KI89A clade	0.01	1.01	1.75	3.32	26.99
Cyanobacteria, Oxyphotobacteria, Chloroplast	0.03	1	0.77	3.29	30.27
Proteobacteria, Gammaproteobacteria, Alteromonadales, Pseudoalteromonadaceae, Pseudoalteromonas	0.01	0.81	2.03	2.67	32.95
Bacteroidetes, Bacteroidia, Chitinophagales, Saprospiraceae	0.01	0.75	1.46	2.47	35.42
Bacteroidetes, Bacteroidia, Chitinophagales	0.01	0.71	1.67	2.35	37.77
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.11	0.63	0.83	2.07	39.83
Proteobacteria, Alphaproteobacteria, Micavibrionales, uncultured, uncultured bacterium, uncultured bacterium	0.01	0.62	1.46	2.05	41.89
Bacteroidetes, Bacteroidia, Chitinophagales, uncultured	0.01	0.59	3.54	1.96	43.84
Planctomycetes, OM190, uncultured bacterium, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.01	0.49	1.73	1.61	45.46
Proteobacteria, Gammaproteobacteria, Cellvibrionales	0.01	0.48	0.94	1.59	47.04
Chloroflexi, Anaerolineae	0.01	0.46	2.25	1.53	48.57
Proteobacteria, Deltaproteobacteria	0.01	0.46	1.07	1.5	50.08
Group T					
Less than 2 samples in group					

Table 3.6 *Similarity Percentage (SIMPER) assessment of the contributions of phylotypes on brood outcomes in the RW in 2021*

SIMPER: Brood Outcomes in the Raceway in 2021					
Phylotypes	**Av.Abund**	**Av.Sim**	**Sim/SD**	**Contrib%**	**Cum.%**
Group Absent					
Average similarity: 28.10					
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.11	6.53	1.41	23.24	23.24
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.04	2.06	0.8	7.34	30.58
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.05	1.66	0.82	5.9	36.48
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.06	1.21	0.5	4.31	40.79
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	1.1	0.76	3.93	44.72
Proteobacteria, Alphaproteobacteria, Rhizobiales, Rhizobiaceae, Hoeflea, Ambiguous taxa	0.02	0.84	0.76	2.99	47.71
Bacteroidetes, Bacteroidia, Chitinophagales	0.13	0.8	1.03	2.84	50.55
Group H_T					
Average similarity: 18.18					
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.2	3.4	SD=0!	18.69	18.69
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.09	2.44	SD=0!	13.45	32.14
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.06	1.14	SD=0!	6.28	38.42
Proteobacteria, Gammaproteobacteria, Vibrionales, Vibrionaceae, Vibrio	0.01	0.73	SD=0!	4.03	42.45
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.02	0.72	SD=0!	3.95	46.4
Actinobacteria, Actinobacteria, Corynebacteriales, Nocardiaceae, Nocardia	0.01	0.63	SD=0!	3.46	49.86
Planctomycetes, OM190, uncultured bacterium, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.01	0.52	SD=0!	2.84	52.71
Group H					
Average similarity: 30.98					
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.1	3.69	0.67	11.92	11.92
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.07	2.51	0.74	8.11	20.02
Bacteroidetes, Bacteroidia, Chitinophagales	0.08	1.89	0.46	6.09	26.12

Table 3.6 (continued).

Proteobacteria, Gammaproteobacteria, Betaproteobacteriales, Burkholderiaceae, Burkholderia-Caballeronia-Paraburkholderia, uncultured beta proteobacterium	0.05	1.49	0.42	4.8	30.92
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.03	1.27	1.16	4.11	35.03
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	1.15	0.59	3.71	38.75
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.03	0.92	0.56	2.96	41.71
Proteobacteria, Gammaproteobacteria	0.02	0.83	1.32	2.67	44.38
Actinobacteria, Actinobacteria, Corynebacteriales, Nocardiaceae, Rhodococcus	0.03	0.72	0.42	2.32	46.7
Planctomycetes, OM190, uncultured bacterium, uncultured bacterium, uncultured bacterium, uncultured bacterium	0.01	0.61	1.16	1.98	48.68
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.02	0.61	0.6	1.97	50.66
Group T					
Less than 2 samples in group					

Table 3.7 *Similarity Percentage (SIMPER) assessment on observed differences in the relative abundances of phylotypes between brood outcomes in the RW in 2021*

SIMPER Pairwise Test: Brood Outcomes in the Raceway in 2021						
Phylotypes	**Av.Abund**	**Av.Abund**	**Av.Diss**	**Diss/SD**	**Contrib%**	**Cum.%**
Groups Absent & H_T						
Average dissimilarity = 68.31	*Group Absent*	*Group H_T*				
Bacteroidetes, Bacteroidia, Chitinophagales	0.13	0.23	13.26	1.08	19.41	19.41
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.05	0.2	9.1	1.16	13.32	32.73

Table 3.7 (continued).

Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.11	0.09	3.88	1.22	5.68	38.4
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.06	0.06	3.7	1.11	5.42	43.82
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira	0.04	0	2.1	0.45	3.07	46.89
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.04	0.01	1.82	1.34	2.66	49.55
Proteobacteria, Gammaproteobacteria, Betaproteobacteriales, Burkholderiaceae, Burkholderia-Caballeronia-Paraburkholderia, uncultured beta proteobacterium	0.01	0.03	1.44	1.08	2.11	51.66
Groups Absent & H						
Average dissimilarity = 69.56	*Group Absent*	*Group H*				
Bacteroidetes, Bacteroidia, Chitinophagales	0.13	0.08	8.18	0.68	11.76	11.76
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.11	0.1	5.41	1.27	7.78	19.54
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.04	0.07	3.56	0.64	5.12	24.66
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.06	0.02	3.11	0.81	4.48	29.13
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira	0.04	0.02	2.65	0.56	3.81	32.94
Proteobacteria, Gammaproteobacteria, Betaproteobacteriales, Burkholderiaceae, Burkholderia-Caballeronia-Paraburkholderia, uncultured beta proteobacterium	0.01	0.05	2.53	0.85	3.64	36.59
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.05	0.03	2.5	1.02	3.59	40.18
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.03	0.03	1.75	1.07	2.52	42.7
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.01	0.03	1.32	0.56	1.9	44.59

Table 3.7 (continued).

Actinobacteria, Actinobacteria, Corynebacteriales, Nocardiaceae, Rhodococcus	0.01	0.03	1.3	0.83	1.86	46.46
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Pseudomonadaceae, Pseudomonas	0.02	0	1.08	0.58	1.56	48.02
Proteobacteria, Alphaproteobacteria, Rhizobiales, Rhizobiaceae, Hoeflea, Ambiguous taxa	0.02	0.01	0.74	1.35	1.06	49.08
Proteobacteria, Alphaproteobacteria, Sphingomonadales, Sphingomonadaceae, Erythrobacter	0.02	0.01	0.73	0.7	1.05	50.13
Groups Absent & T						
Average dissimilarity = 91.08	*Group Absent*	*Group T*				
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.01	0.75	37.28	116.04	40.94	40.94
Bacteroidetes, Bacteroidia, Chitinophagales	0.13	0.01	6.13	0.44	6.73	47.66
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.11	0	5.47	1.39	6.01	53.67
Groups H_T & H						
Average dissimilarity = 71.13	*Group H_T*	*Group H*				
Bacteroidetes, Bacteroidia, Chitinophagales	0.23	0.08	11.19	1.14	15.74	15.74
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Spongiibacteraceae	0.2	0.03	9.42	1.15	13.24	28.98
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.09	0.1	4.99	1.14	7.02	35.99
Proteobacteria, Alphaproteobacteria, Rhodobacterales, Rhodobacteraceae	0.01	0.07	3.27	0.54	4.59	40.58
Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxellaceae	0.06	0.02	2.76	1.22	3.88	44.47
Proteobacteria, Gammaproteobacteria, Betaproteobacteriales, Burkholderiaceae, Burkholderia-Caballeronia-Paraburkholderia, uncultured beta proteobacterium	0.03	0.05	2.71	1.07	3.81	48.28
Bacteroidetes, Bacteroidia, Flavobacteriales, Flavobacteriaceae	0.02	0.03	1.62	0.99	2.28	50.56

Table 3.7 (continued).

Groups H_T & T						
Average dissimilarity = 93.91	*Group H_T*	*Group T*				
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0	0.75	37.46	266.28	39.89	39.89
Bacteroidetes, Bacteroidia, Chitinophagales	0.23	0.01	11.12	0.74	11.84	51.73
Groups H & T						
Average dissimilarity = 88.66	*Group H*	*Group T*				
Proteobacteria, Gammaproteobacteria, Alteromonadales, Alteromonadaceae, Alteromonas	0.03	0.75	36.08	14.78	40.69	40.69
Proteobacteria, Gammaproteobacteria, Cellvibrionales, Cellvibrionaceae	0.1	0	4.73	0.8	5.34	46.02
Bacteroidetes, Bacteroidia, Chitinophagales	0.08	0.01	3.46	0.55	3.9	49.92
Proteobacteria, Alphaproteobacteria, Rhodospirillales, Thalassospiraceae, Thalassospira	0.02	0.07	3.36	2.06	3.79	53.71

The RW dataset was visualized via an nMDS plot (Figure 3.2) to determine orientation of samples to each other based on survival rates and the relative abundances of nine phylotypes. Phylotypes were selected based on the results of the SIMPER analysis (Tables 3.3-3.7), and ultimately, members of the families Rhodobacteraceae, Alteromonadaceae, Flavobacteriaceae, Saccharospirillaceae, and Cellvibrionaceae as well as a member of the order Chitinophagales and chloroplast were chosen for visualization. Samples mostly associated with low performing broods (survival rates of <10%) had high abundances of the members from families Rhodobacteraceae, Alteromonadaceae, and Flavobacteriaceae as well as the member of order Chitinophagales and chloroplast. Samples associated with high performing larvae (>15%) had high abundances of primarily the member of family Cellvibrionaceae and lesser abundances of Rhodobacteraceae, Alteromonadaceae, and one chloroplast phylotype.

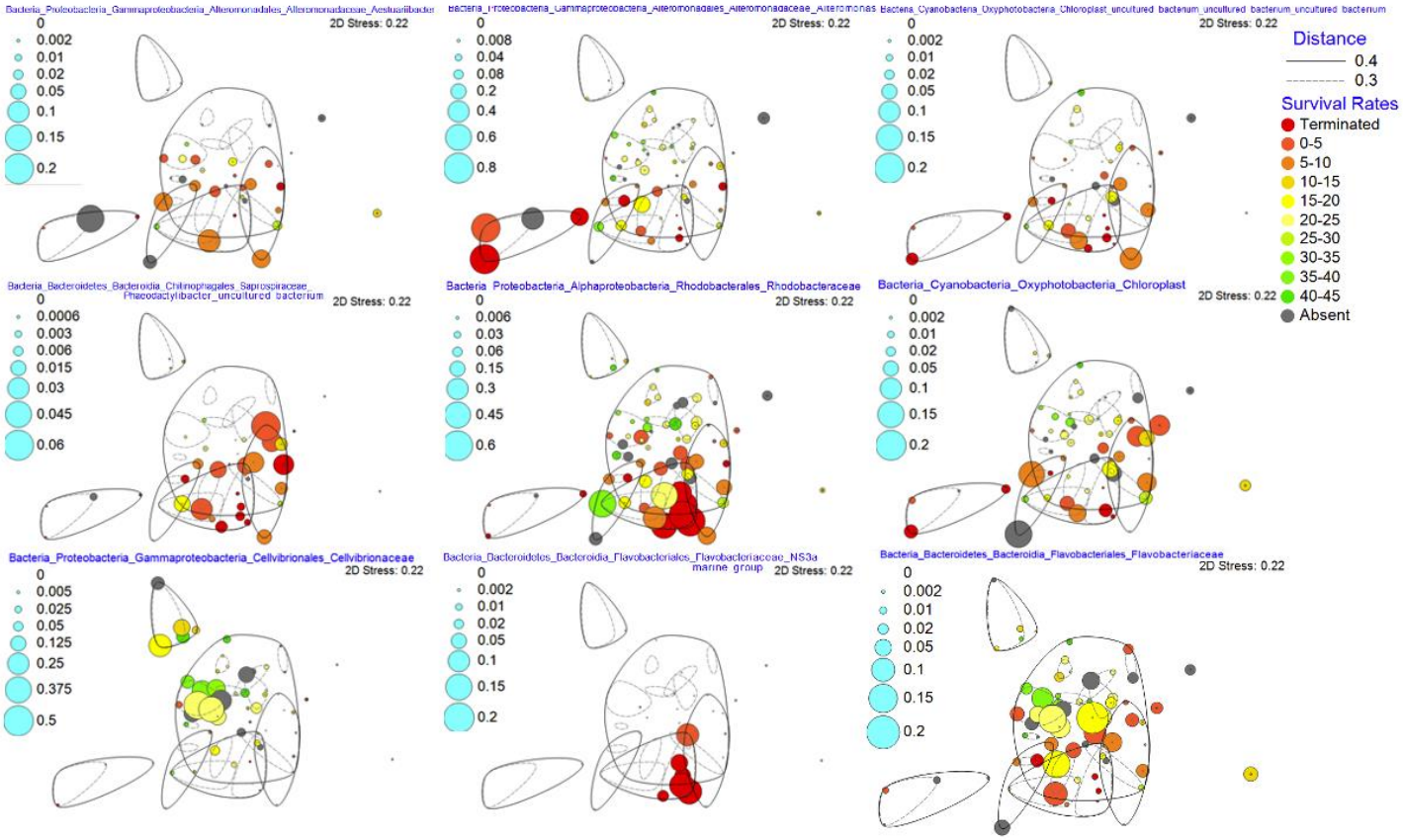

Figure 3.2 *Non-metric multidimensional scaling analysis (nMDS) for the RW across runs weighted by relative abundance*

Data derived from the **weighted UniFrac distance matrix and relative abundance data. Samples are grouped by categories of larval survival rates for visualization. Categories of survival define Terminated, 0-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-30%, 30-35%, 35-40%, 40-45%, and the absence of larvae. Symbols are weighted by relative abundance of prominent phylotypes contributing to brood outcomes in the RW. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.3 (70% similar).**

3.3.3 T Analysis

The T dataset was visualized via an nMDS plot (Figure 3.3) to determine orientation of samples to each other based on brood outcome (left) and survival rates (right). In 2019, samples taken during broods that were harvested generally oriented above other 2019 samples. Samples taken during broods that were terminated generally oriented below other 2019 samples. Samples taken when broods of different outcomes overlapped mostly occupied the middle right side of the plot between harvested and terminated samples. Samples taken when larvae were absent were mostly outlier samples. In 2020, the only sample types were harvested and absent since no broods overlapped and no T samples taken during the single terminated brood's rearing were included in the dataset. Harvested and absent samples from 2020 mostly group together in the top of a 0.35 distance (65% similarity) cluster with 2021 samples. One absent sample is an outlier. In 2021, samples associated with all brood outcomes mostly oriented together. Some harvested samples grouped with 2019 harvested and overlapping samples. Overall, grouping in the T appears to be based on run more than on brood outcome.

Further visualizing T samples by larval survival rates, samples taken during low performing broods (survival rates <10%) mostly grouped together near terminated samples. Samples associated with higher performing broods (>30%) mostly orient in the middle of the plot. Samples associated with survival rates between 10% and 30% are spread across the plot near low performing, high performing, and terminated samples. Though there may be some grouping by samples associated with high and low/terminated broods, samples are primarily organized by run rather than larval survival rates. 2019 samples being spread across the right side of the plot with some differentiation between

terminated and harvested broods is likely explained by previously investigated confounders in 2019 that caused larval mortalities (see chapter 2.3.11).

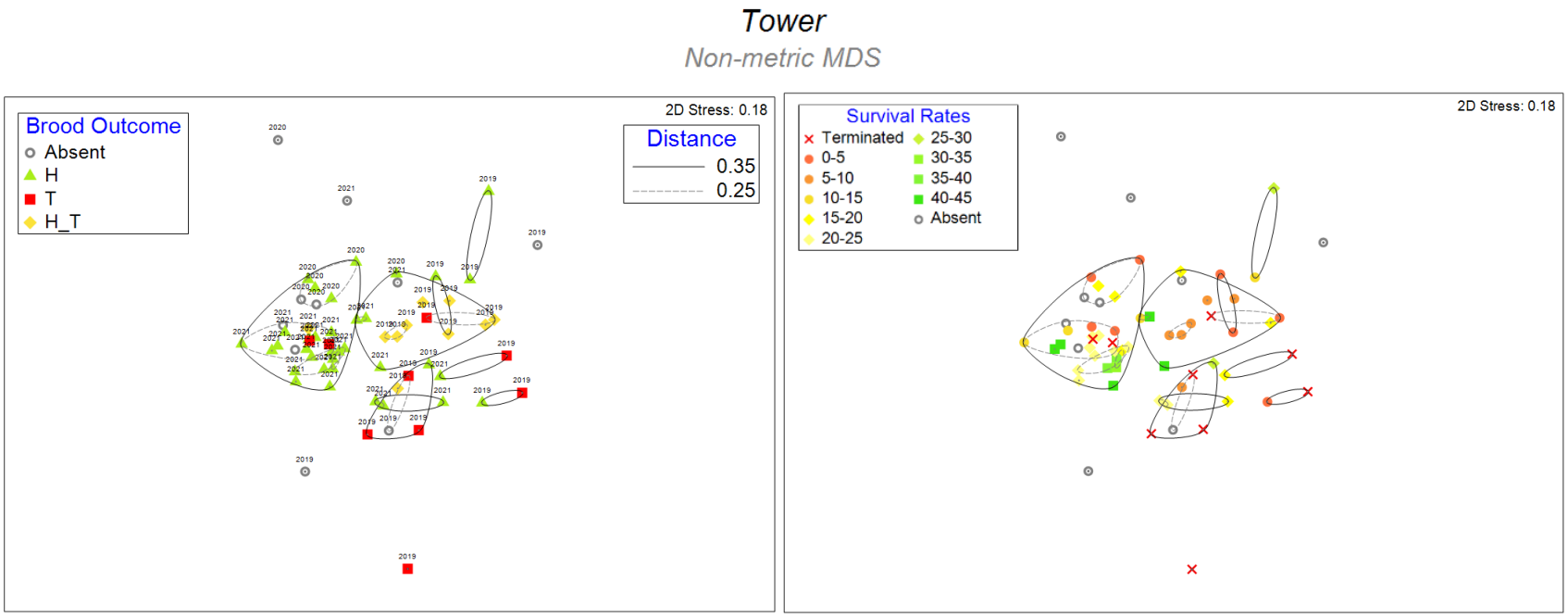

Figure 3.3 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the T during all three runs visualizing larval performance*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by brood outcomes (left) and percent survival (right) for visualization. Outcomes consist of harvested (H), terminated, overlap of harvested and terminated broods (H_T), and absence of larvae. Numbers above each symbol indicate the run the sample was taken. Categories of survival define Terminated, 0-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-30%, 30-35%, 35-40%, 40-45%, and the absence of larvae. A CLUSTER analysis generated distance contour lines at 0.35 (65% similar) and 0.25 (75% similar).

A BEST procedure was performed on larval survival rates and brood outcomes in the T in each run to determine the degree to which larval performance shapes the microbiome (Table 3.8). Across all three runs, no single or combined variables provided a spearman rank correlation coefficient that exceeded 0.000. Compared to water quality parameters in the T (see chapter 2.3.6), larval performance in terms of survival rates and brood outcomes did not have as strong an influence on the T microbiome across all three runs. Comparing results across all three runs, survival rates had the highest correlation in 2020 and the lowest in 2019 while brood outcome had the highest correlation in 2019 and lowest in 2021. These results suggest little to no relationship between larval performance and the composition of the T microbiome in all three runs.

Table 3.8 *BEST procedure matching the T's weighted UniFrac distance matrix in correlation with larval survival rates and brood outcomes*

BEST Results: Larval Performance in the Tower								
Correlation method: Spearman rank								
Maximum number of variables: 2								
Analyze between: Samples from T Weighted UniFrac Distance Matrix								
Resemblance measure: D1 Euclidean distance								
Run 2019			*Run 2020*			*Run 2021*		
Number of variables: 1			*Number of variables: 1*			*Number of variables: 1*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	-0.036	Brood Outcome	1	-0.151	% Survival	1	-0.036	Brood Outcome
1	-0.175	% Survival				1	-0.175	% Survival
Number of variables: 2			*Number of variables: 2*			*Number of variables: 2*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
2	-0.172	% Survival, Brood Outcome	2	-0.151	% Survival, Brood Outcome	2	-0.172	% Survival, Brood Outcome
Best result for each number of variables			*Best result for each number of variables*			*Best result for each number of variables*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	-0.036	Brood Outcome	1	-0.151	% Survival	1	-0.036	Brood Outcome
2	-0.172	% Survival, Brood Outcome	2	-0.151	% Survival, Brood Outcome	2	-0.172	% Survival, Brood Outcome
Best results			*Best results*			*Best results*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	-0.036	Brood Outcome	1	-0.151	% Survival	1	-0.036	Brood Outcome
2	-0.172	% Survival, Brood Outcome	2	-0.151	% Survival, Brood Outcome	2	-0.172	% Survival, Brood Outcome
1	-0.175	% Survival				1	-0.175	% Survival

3.3.4 LT Analysis

The LT dataset was visualized via an nMDS plot (Figure 3.4) to determine orientation of samples to each other based on brood outcome (left) and survival rates

(right). In 2019, samples taken during broods that were harvested grouped loosely with other harvested samples from 2021. Samples taken during broods that were terminated grouped tightly within two 0.2 distance (80% similarity) clusters in the middle of the plot. No 2019 samples in this dataset were taken when larvae were absent. All samples in 2019 were labeled according to which brood existed in the same larval tank as the associated sequencing samples were taken, so there are no overlapping outcome samples. Harvested samples from 2020 are mostly grouped in a 0.2 distance cluster. The single terminated sample from 2020 is grouped with other terminated samples. No broods overlapped in 2020. The single 2020 sample taken when larvae were absent is positioned near a harvested 2021 sample. Harvested samples in 2021 are spread across the plot. One of the two terminated samples are grouped with other terminated samples from 2019 and 2020 while the other sample is grouped with a harvested 2021 sample. The single overlapping sample in 2021 is oriented with other harvested samples. Samples taken when larvae were absent in 2021 orient with harvested 2021 samples.

Further visualizing LT samples by larval survival rates, samples associated with low (<10%), medium (10-30%), and high (>30%) performing larval broods are spread across the plot with no clear grouping by performance. Terminated samples mostly orient with other terminated samples. There is no clear pattern based on larval survival rates in the LT.

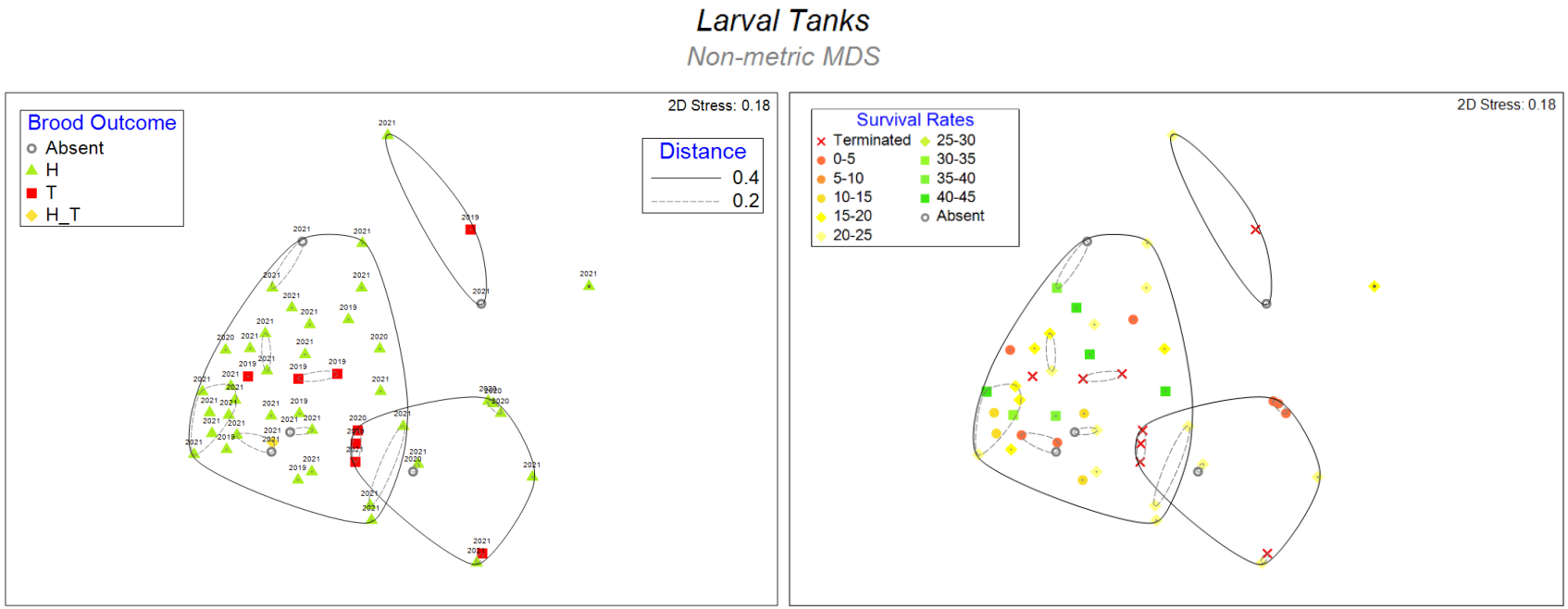

Figure 3.4 *Non-metric multidimensional scaling analysis (nMDS) on samples collected from the LT during all three runs visualizing larval performance*

Data derived from the weighted UniFrac distance matrix. Samples are grouped by brood outcomes (left) and percent survival (right) for visualization. Outcomes consist of harvested (H), terminated, overlap of harvested and terminated broods (H_T), and absence of larvae. Numbers above each symbol indicate the run the sample was taken. Categories of survival define Terminated, 0-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-30%, 30-35%, 35-40%, 40-45%, and the absence of larvae. A CLUSTER analysis generated distance contour lines at 0.4 (60% similar) and 0.2 (80% similar)

A BEST procedure was performed on larval survival rates and brood outcomes in the LT in each run to determine the degree to which larval performance shapes the microbiome (Table 27). In 2019, survival rates had a higher individual correlation with community composition compared to brood outcomes (0.051 and 0.015, respectively). A combination of variables cumulated in a correlation of 0.054. In 2020, no single or combined variables provided a spearman rank correlation coefficient that exceeded 0.000. In 2021, brood outcomes had a higher individual correlation with community composition compared to survival rates (0.037 and -0.062, respectively). A combination of variables cumulated in a correlation of -0.062. The LT has no similar analysis involving water quality parameters to compare to. Comparing the LT results to that of the RW and T, larval performance has a stronger influence on the LT microbiome than it does on the T microbiome but a lesser influence than on the RW microbiome. Comparing results across all three runs, survival rates had the highest correlation in 2019 and the lowest in 2021 while brood outcome had the highest correlation in 2021 and lowest in 2020. These results suggest a relationship does exist between larval performance and the composition of the LT microbiome mostly in 2019 and 2021.

Table 3.9 *BEST procedure matching the LT's weighted UniFrac distance matrix in correlation with larval survival rates and brood outcomes*

BEST Results: Larval Performance in the Larval Tanks								
Correlation method: Spearman rank								
Maximum number of variables: 2								
Analyze between: Samples from LT Weighted UniFrac Distance Matrix								
Resemblance measure: D1 Euclidean distance								
Run 2019			*Run 2020*			*Run 2021*		
Number of variables: 1			*Number of variables: 1*			*Number of variables: 1*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.051	% Survival	1	-0.057	% Survival	1	0.037	Brood Outcome
1	0.015	Brood Outcome	1	-0.074	Brood Outcome	1	-0.062	% Survival
Number of variables: 2			*Number of variables: 2*			*Number of variables: 2*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
2	0.054	% Survival, Brood Outcome	2	-0.057	% Survival, Brood Outcome	2	-0.062	% Survival, Brood Outcome
Best result for each number of variables			*Best result for each number of variables*			*Best result for each number of variables*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
1	0.051	% Survival	1	-0.057	% Survival	1	0.037	Brood Outcome
2	0.054	% Survival, Brood Outcome	2	-0.057	% Survival, Brood Outcome	2	-0.062	% Survival, Brood Outcome
Best results			*Best results*			*Best results*		
No.Vars	Corr.	Selections	No.Vars	Corr.	Selections	No.Vars	Corr.	Selections
2	0.054	% Survival, Brood Outcome	1	-0.057	% Survival	1	0.037	Brood Outcome
1	0.051	% Survival	2	-0.057	% Survival, Brood Outcome	1	-0.062	% Survival
1	0.015	Brood Outcome	1	-0.074	Brood Outcome	2	-0.062	% Survival, Brood Outcome

3.4 Discussion

3.4.1 Larval Health Connection to the RAS Microbiome

There has been a focus on researching RAS microbiome dynamics (Attramadal *et al.*, 2012, Attramadal *et al.*, 2014, Bakke *et al.*, 2017, Dahle *et al.*, 2022) and oyster larval performance in aquaculture settings (Kuhn *et al.*, 2013, Chen *et al.*, 2016, Qiu *et al.*, 2017) as separate entities. Few studies observe features of both aquaculture and oyster larval microbiomes (Asmani *et al.*, 2016, Arfken *et al.*, 2021), and this study is the first to determine a possible link between oyster larval survival and RAS microbial stability. Additionally, contributions of specific phylotypes to production outcomes were investigated with the sub-goal of identifying potential indicators of larval health in the RAS microbiome.

Oyster larvae rely on their environment to, in part, source their core microbiomes (Asmani *et al.*, 2016, Laroche *et al.*, 2018). Even as adults when their microbiome is specialized for specific tissues (King *et al.*, 2012, Pierce *et al.*, 2016, Pimentel *et al.*, 2021), their microbial composition reflects the composition of the water microbiome they inhabited as larvae (Arfken *et al.*, 2021). Previous works have observed the water microbiome associated with oyster larvae in hatcheries is made up predominately of Alphaproteobacteria, Flavobacteria, and Chlamydie (Stevick *et al.*, 2019, Arfken *et al.*, 2021). The core microbiome of *C. virginica* larvae, specifically, has been defined primarily by the families Rhodobacteraceae, Flavobacteriaceae, and Alteromonadaceae of the phyla Proteobacteria, Bacteroidota, and Pseudomonadota, respectively. The present study found the RAS microbiome is consistent with these findings, namely in the larval tanks (see Figure 2.5). The RW and T have diverse microbiomes populated by members

of classes Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia. The LT microbiome is sparser, being comprised mostly of families Rhodobacteraceae, Alteromonadaceae, Flavobacteriaceae, and Pseudomonadaceae. Compared to the RW and T, the LT microbiome more closely resembles the larval core microbiome.

Characterizing the RAS microbiomes is important to understanding the way in which larval survival influences community composition (or vice versa) both in the water the larvae inhabit (LT) and the water reclamation stages (RW and T). This study found brood outcomes (harvested or terminated) and survival rates shaped the RW microbiome but not the T or LT microbiomes. In the RW, the compartment directly receiving water from the LT, terminated broods were associated with a community composition distinct from the composition associated with harvested broods. Signature communities emerged in the RW microbiome during mortalities. These communities were distinguished from periods of higher larval survival rates by an abundance of members from Rhodobacteraceae, Alteromonadaceae, Thalassospiraceae, Flavobacteriaceae, Saccharospirillaceae, and Cellvibrionaceae as well as chloroplast from the class Oxyphotobacteria, as evidenced by the SIMPER analyses performed on the RW microbiome (Tables 3.3-3.7). This is in contrast to King *et al.* (2019) where mortality in *C. gigas* was primarily associated with members of Bacteroides and specifically Pseudoalteromonadaceae. The prominent families noted in the present study were, however, members of the larval microbiome and diet (chloroplast). Rhodobacteraceae is the dominant family in the larval core microbiome, and members of the family populated the entire system. Flavobacteriaceae is the second most abundant family in the core microbiome (as defined by Arfken *et al.*, 2021) and, like Rhodobacteraceae, is found

throughout the RAS. These taxa, in particular, are found in elevated abundances in the RW microbiome during larval mortalities.

No statistical changes in community composition were observed in the T or LT microbiomes. A lack or heavily muted response to larval activities in the T is likely due to the distance between the two compartments. Water in the T undergoes two prior rounds of filtration before entering the compartment. It was expected the compartment with the most direct contact with the larvae, the LTs, would have community shifts associated with larval health. Direct contact, however, may have been the reason why no shifts were detected. Across three years in this experiment, the most successful brood only reached a final percent survival of about 44%. Larval deaths are a feature of the LT, hence taxa associated with the core larval microbiome are continuously found in higher relative abundances in the LT microbiome compared to the RW or T. Larval mortality's impact on RAS microbiomes are not ignored by the LT but, rather, are louder in the RW.

Patterns in microbial stability in the RW and LT reflect changes in brood performances, indicating larval health and the RAS microbiome are connected via the rearing tanks and the first stage of water reclamation. The moving window analyses performed in chapter 2 detailed the weighted UniFrac distances between successive samples, providing a means of assessing stability over time. It was previously noted patterns of stability and instability corresponded to transitions between broods, though they did not match exactly. With the added assessment of larval survival, periods of instability overlap with transitions from low to high performing broods and vice versa. For example, a transition from 2021's brood 6 to brood 7 corresponded to the highest moment of instability in the RW microbiome. Brood 6 was terminated for low larval

health while brood 7 was the aforementioned highest performing brood across all three years. The subsequent transition from brood 7 to brood 8 (18% survival) was marked by the highest moment of instability in the LT microbiome. Conversely, transitions from similarly performing broods, such as 2021's broods 3 and 4, were marked by periods of stability in both the RW and LT microbiomes. Previously identified instability associated with confounders in 2019 (chapter 2.3.11) likewise display declined stability during the mortality event and a subsequent stabilization period after the mortalities ended. There is a clear link between RAS microbiome stability declining alongside declines in larval performance. It may be advantageous to avoid the period of instability by allowing the microbiomes to recover before introducing a new brood. Further research into the connections between mortalities and microbial instability is recommended.

3.5 Conclusions

There are communities associated with low larval survival and poor brood outcomes comprised of elevated abundances of the families Rhodobacteraceae, Alteromonadaceae, Flavobacteriaceae, and members of class Oxyphotobacteria. This study provides a first look at RAS communities indicative of production losses. It is clear there is a link between microbiome stability and larval performance that may be informative for RAS operations. New broods that were introduced to the system immediately after prior brood(s) performed poorly aligned with periods of instability in the RW and LT, and the microbiomes stabilized the longer the brood survived in the system.

REFERENCES

Ahmed N & Turchini GM (2021) Recirculating aquaculture systems (RAS): Environmental solution and climate change adaptation. Journal of Cleaner Production 297: 126604.

Allison SD & Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* **105**: 11512-11519.

Almeida DB, Magalhães C, Sousa Z, Borges MT, Silva E, Blanquet I & Mucha AP (2021) Microbial community dynamics in a hatchery recirculating aquaculture system (RAS) of sole (Solea senegalensis). *Aquaculture* **539**: 736592.

Arfken A, Song B, Allen SK & Carnegie RB (2021) Comparing larval microbiomes of the eastern oyster (Crassostrea virginica) raised in different hatcheries. *Aquaculture* **531**: 735955.

Asmani K, Petton B, Le Grand J, Mounier J, Robert R & Nicolas J-L (2016) Establishment of microbiota in larval culture of Pacific oyster, Crassostrea gigas. *Aquaculture* **464**: 434-444.

Asmani K, Petton B, Le Grand J, Mounier J, Robert R & Nicolas J-L (2017) Determination of stocking density limits for Crassostrea gigas larvae reared in flow-through and recirculating aquaculture systems and interaction between larval density and biofilm formation. *Aquatic Living Resources* **30**: 29.

Attramadal KJK, Truong TMH, Bakke I, Skjermo J, Olsen Y & Vadstein O (2014) RAS and microbial maturation as tools for K-selection of microbial communities improve survival in cod larvae. *Aquaculture* **432**: 483-490.

Attramadal KJK, Salvesen I, Xue R, Øie G, Størseth TR, Vadstein O & Olsen Y (2012) Recirculation as a possible microbial control strategy in the production of marine larvae. *Aquacultural Engineering* **46**: 27-39.

Attramadal KJK, Øien JV, Kristensen E, Evjemo JO, Kjørsvik E, Vadstein O & Bakke I (2021) UV treatment in RAS influences the rearing water microbiota and reduces the survival of European lobster larvae (Homarus gammarus). *Aquacultural Engineering* **94**: 102176.

Badiola M, Mendiola D & Bostock J (2012) Recirculating Aquaculture Systems (RAS) analysis: Main issues on management and future challenges. *Aquacultural Engineering* **51**: 26-35.

Bakke I, Åm AL, Kolarevic J, Ytrestøyl T, Vadstein O, Attramadal KJK & Terjesen BF (2017) Microbial community dynamics in semi-commercial RAS for production of Atlantic salmon post-smolts at different salinities. *Aquacultural Engineering* **78**: 42-49.

Balami S (2019) Recirculation Aquaculture Systems: Components, Advantages, And Drawbacks. *Education* **2021**.

Bartelme RP, Smith MC, Sepulveda-Villet OJ & Newton RJ (2019) Component Microenvironments and System Biogeography Structure Microorganism Distributions in Recirculating Aquaculture and Aquaponic Systems. *mSphere* **4**: e00143-00119.

Blancheton JP, Attramadal KJK, Michaud L, d'Orbcastel ER & Vadstein O (2013) Insight into bacterial population in aquaculture systems and its implication. *Aquacultural Engineering* **53**: 30-39.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M & Asnicar F (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology* **37**: 852-857.

Bukin YS, Galachyants YP, Morozov IV, Bukin SV, Zakharenko AS & Zemskaya TI (2019) The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific Data* **6**: 190007.

Calabrese A, Collier RS, Nelson DA & MacInnes JR (1973) The toxicity of heavy metals to embryos of the American oyster Crassostrea virginica. *Marine Biology* **18**: 162-166.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA & Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods* **13**: 581-583.

Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM & Young VB (2008) Decreased Diversity of the Fecal Microbiome in Recurrent Clostridium difficile—Associated Diarrhea. *The Journal of Infectious Diseases* **197**: 435-438.

Chen H, Liu Z, Shi Y & Ding H (2016) Microbiological analysis and microbiota in oyster: a review. *Invertebrate Survival Journal* **13**: 374-388.

Clarke KR & Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. *Marine Ecology Progress Series* **92**: 205-219.

Comeau AM, Li WK, Tremblay J-É, Carmack EC & Lovejoy C (2011) Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum. *PloS one* **6**: e27492.

Dahle SW, Attramadal KJK, Vadstein O, Hestdahl HI & Bakke I (2022) Microbial community dynamics in a commercial RAS for production of Atlantic salmon fry (Salmo salar). *Aquaculture* **546**: 737382.

Dahle SW, Bakke I, Birkeland M, Nordøy K, Dalum AS & Attramadal KJK (2020) Production of lumpfish (Cyclopterus lumpus L.) in RAS with distinct water treatments: Effects on fish survival, growth, gill health and microbial communities in rearing water and biofilm. *Aquaculture* **522**: 735097.

Dalsgaard J, Lund I, Thorarinsdottir R, Drengstig A, Arvonen K & Pedersen PB (2013) Farming different species in RAS in Nordic countries: Current status and future perspectives. *Aquacultural Engineering* **53**: 2-13.

Ebeling JM & Timmons MB (2012) Recirculating Aquaculture Systems. *Aquaculture Production Systems*,p.^pp. 245-277.

Fornshell G, Hinshaw J & Tidwell JH (2012) Flow-through Raceways. *Aquaculture Production Systems*,p.^pp. 173-190.

Gray MW, Alexander ST, Beal BF, *et al.* (2022) Hatchery crashes among shellfish research hatcheries along the Atlantic coast of the United States: A case study of production analysis at Horn Point Laboratory. *Aquaculture* **546**: 737259.

Green TJ & Barnes AC (2010) Bacterial diversity of the digestive gland of Sydney rock oysters, Saccostrea glomerata infected with the paramyxean parasite, Marteilia sydneyi. *Journal of Applied Microbiology* **109**: 613-622.

Green TJ, Siboni N, King WL, Labbate M, Seymour JR & Raftos D (2019) Simulated Marine Heat Wave Alters Abundance and Structure of Vibrio Populations Associated with the Pacific Oyster Resulting in a Mass Mortality Event. *Microbial Ecology* **77**: 736-747.

Hamdan LJ, Coffin RB, Sikaroodi M, Greinert J, Treude T & Gillevet PM (2013) Ocean currents shape the microbiome of Arctic marine sediments. *The ISME Journal* **7**: 685-696.

Hernández-Zárate G & Olmos-Soto J (2006) Identification of bacterial diversity in the oyster Crassostrea gigas by fluorescent in situ hybridization and polymerase chain reaction. *Journal of Applied Microbiology* **100**: 664-672.

Kanisan DP, Quek ZBR, Oh RM, Afiq-Rosli L, Lee JN, Huang D & Wainwright BJ (2023) Diversity and Distribution of Microbial Communities Associated with Reef Corals of the Malay Peninsula. *Microbial Ecology* **85**: 37-48.

Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30**: 772-780.

King GM, Judd C, Kuske CR & Smith C (2012) Analysis of stomach and gut microbiomes of the eastern oyster (Crassostrea virginica) from coastal Louisiana, USA. *PLoS one* **7**: e51475.

Kuhn DD, Angier MW, Barbour SL, Smith SA & Flick GJ (2013) Culture feasibility of eastern oysters (Crassostrea virginica) in zero-water exchange recirculating aquaculture systems using synthetically derived seawater and live feeds. *Aquacultural Engineering* **54**: 45-48.

Laroche O, Symonds JE, Smith KF, Banks JC, Mae H, Bowman JP & Pochon X (2018) Understanding bacterial communities for informed biosecurity and improved larval survival in Pacific oysters. *Aquaculture* **497**: 164-173.

Leys C, Ley C, Klein O, Bernard P & Licata L (2013) Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* **49**: 764-766.

Lokmer A & Wegner K (2015) Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. *The ISME Journal* **9**: 670-682.

Lokmer A, Kuenzel S, Baines JF & Wegner KM (2016) The role of tissue-specific microbiota in initial establishment success of Pacific oysters. *Environmental Microbiology* **18**: 970-987.

Lokmer A, Goedknegt MA, Thieltges DW, Fiorentino D, Kuenzel S, Baines JF & Wegner KM (2016) Spatial and temporal dynamics of Pacific oyster hemolymph microbiota across multiple scales. *Frontiers in microbiology* **7**: 1367.

Martin M (2011) CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**.

Martins CIM, Eding EH, Verdegem MCJ, Heinsbroek LTN, Schneider O, Blancheton JP, d'Orbcastel ER & Verreth JAJ (2010) New developments in recirculating aquaculture systems in Europe: A perspective on environmental sustainability. *Aquacultural Engineering* **43**: 83-93.

Meyer M & Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* **2010**: pdb. prot5448.

Midilli A, Kucuk H & Dincer I (2012) Environmental and sustainability aspects of a recirculating aquaculture system. *Environmental Progress & Sustainable Energy* **31**: 604-611.

Mohamed Ramli N, Giatsis C, Md Yusoff F, Verreth J & Verdegem M (2018) Resistance and resilience of small-scale recirculating aquaculture systems (RAS) with or without algae to pH perturbation. *PloS one* **13**: e0195862.

Offret C, Paulino S, Gauthier O*, et al.* (2020) The marine intertidal zone shapes oyster and clam digestive bacterial microbiota. *FEMS Microbiology Ecology* **96**.

Paillard C, Gueguen Y, Wegner KM, Bass D, Pallavicini A, Vezzulli L & Arzul I (2022) Recent advances in bivalve-microbiota interactions for disease prevention in aquaculture. *Current Opinion in Biotechnology* **73**: 225-232.

Parada AE, Needham DM & Fuhrman JA (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**: 1403-1414.

Pierce ML & Ward JE (2018) Microbial Ecology of the Bivalvia, with an Emphasis on the Family Ostreidae. *Journal of Shellfish Research* **37**: 793-806, 714.

Pierce ML & Ward JE (2019) Gut Microbiomes of the Eastern Oyster (*Crassostrea virginica*) and the Blue Mussel (*Mytilus edulis*): Temporal Variation and the Influence of Marine Aggregate-Associated Microbial Communities. *mSphere* **4**: e00730-00719.

Pierce ML, Ward JE, Holohan BA, Zhao X & Hicks RE (2016) The influence of site and season on the gut and pallial fluid microbial communities of the eastern oyster,

Crassostrea virginica (Bivalvia, Ostreidae): community-level physiological profiling and genetic structure. *Hydrobiologia* **765**: 97-113.

Pillai P, Gouhier TC & Vollmer SV (2014) The cryptic role of biodiversity in the emergence of host–microbial mutualisms. *Ecology Letters* **17**: 1437-1446.

Pimentel ZT, Dufault-Thompson K, Russo KT, Scro AK, Smolowitz RM, Gomez-Chiarri M & Zhang Y (2021) Microbiome Analysis Reveals Diversity and Function of Mollicutes Associated with the Eastern Oyster, Crassostrea virginica. *mSphere* **6**: e00227-00221.

Prado S, Romalde JL & Barja JL (2010) Review of probiotics for use in bivalve hatcheries. *Veterinary Microbiology* **145**: 187-197.

Qiu T, Qi J, Zheng J & Liu Y (2017) Design and performance of a recirculating aquaculture system for oyster larval culture. *Aquaculture Research* **48**: 5699-5706.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**: D590-D596.

Ramos CdO, da Silva FC, Gomes CHAdM, Langdon C, Takano P, Gray MW & de Melo CMR (2021) Effect of larval density on growth and survival of the Pacific oyster Crassostrea gigas in a recirculation aquaculture system. *Aquaculture* **540**: 736667.

Rodionov, S. N. (2016). A comparison of two methods for detecting abrupt changes in the variance of climatic time series. *Advances in Statistical Climatology, Meteorology and Oceanography*, **2(1)**: 63-78.

Rognes T, Flouri T, Nichols B, Quince C & Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**: e2584.

Rud I, Kolarevic J, Holan AB, Berget I, Calabrese S & Terjesen BF (2017) Deep-sequencing of the bacterial microbiota in commercial-scale recirculating and semi-closed aquaculture systems for Atlantic salmon post-smolt production. *Aquacultural Engineering* **78**: 50-62.

Simons AL, Churches N & Nuzhdin S (2018) High turnover of faecal microbiome from algal feedstock experimental manipulations in the Pacific oyster (Crassostrea gigas). *Microbial Biotechnology* **11**: 848-858.

Stevick RJ, Sohn S, Modak TH, Nelson DR, Rowley DC, Tammi K, Smolowitz R, Markey Lundgren K, Post AF & Gómez-Chiarri M (2019) Bacterial community dynamics in an oyster hatchery in response to probiotic treatment. *Frontiers in microbiology* **10**: 1060.

Terjesen BF, Summerfelt ST, Nerland S*, et al.* (2013) Design, dimensioning, and performance of a research facility for studies on the requirements of fish in RAS environments. *Aquacultural Engineering* **54**: 49-63.

Tidwell JH (2012) Functions and Characteristics of All Aquaculture Systems. *Aquaculture Production Systems*,p.^pp. 51-63.

Tidwell JH (2012) Characterization and Categories of Aquaculture Production Systems. *Aquaculture Production Systems*,p.^pp. 64-78.

Trabal Fernández N, Mazón-Suástegui JM, Vázquez-Juárez R, Ascencio-Valle F & Romero J (2014) Changes in the composition and diversity of the bacterial microbiota associated with oysters (Crassostrea corteziensis, Crassostrea gigas

and Crassostrea sikamea) during commercial production. *FEMS Microbiology Ecology* **88**: 69-83.

Tucker C & Hargreaves J (2012) Ponds. *Aquaculture Production Systems*,p.^pp. 191-244.

Utting SD & Helm MM (1985) Improvement of sea water quality by physical and chemical pre-treatment in a bivalve hatchery. *Aquaculture* **44**: 133-144.

Vignier J, Laroche O, Rolton A, Wadsworth P, Kumanan K, Trochel B, Pochon X & King N (2021) Dietary exposure of Pacific oyster (Crassostrea gigas) larvae to compromised microalgae results in impaired fitness and microbiome shift. *Frontiers in microbiology* 2319.

Walters W, Hyde ER, Berg-Lyons D, *et al.* (2016) Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**: e00009-00015.

Wegner KM, Volkenborn N, Peter H & Eiler A (2013) Disturbance induced decoupling between host genetics and composition of the associated microbiome. *BMC Microbiology* **13**: 252.

Weng N & Wang W-X (2014) Improved tolerance of metals in contaminated oyster larvae. *Aquatic Toxicology* **146**: 61-69.