Distribution of microbial eukaryotes living as epibiota on tube worms (*Ridgeia piscesae*) in hydrothermal fluid flow and vent plumes at Axial Seamount

Han Weinrich

University of Washington

School of Oceanography, Box 357940

Seattle WA 98195

hweinric@uw.edu

06/02/2023

Abstract

Hydrothermal vents are deep-sea environments characterized by high temperature, low pH, low oxygen concentration and high chemical concentrations with a chemosynthesis based ecosystem. Microbial eukaryotes (protists) live in the water column and attached to hard substrates in the vent environment, feeding on chemosynthetic prokaryotes and detritus. A proportion of these protists are endemic to the vent environment, while other taxa are cosmopolitan. I hypothesized that protists living attached to tubeworms in hydrothermal flow would possess adaptations that allow them to survive vent conditions, leading to a lower community diversity and species richness in these communities than in the water column above the vent. Paired microbial samples were collected from the exterior of tube worms and the water column at active venting sites at Axial Seamount. Sample DNA was amplified using universal eukaryotic and metazoan-excluding primer sets and the v4 hypervariable region of the 18s gene was sequenced. Non-metazoan primers amplified prokaryotic sequences and comparatively fewer protists than universal primers, while universal primers were able to amplify more protist sequences from worm samples than non-metazoan primers suggesting that they can be effectively implemented to amplify epibiotic sequences. Higher species richness and diversity was observed in the water column samples than tube worm samples using both sets of primers, suggesting higher rates of endemism living on worms in direct hydrothermal flow as hypothesized.

Plain Language Summary

Hydrothermal vent ecosystems are based on the chemicals inside fluids coming out of vents from beneath the seafloor. Specially adapted bacteria and archaea use these chemicals to power their metabolisms and fix carbon (chemosynthesis). Protists at vents are single celled eukaryotic organisms that feed on chemosynthetic bacteria and archaea along with detritus sinking from the upper ocean. Some protists living in and around vents are not unique to vent ecosystems - they can be found elsewhere in the ocean, whereas some are endemic, meaning they are only found at vents. Protists living in and around vent fluid will need to be able to withstand the low pH, high temperatures, low oxygen and toxic chemicals there. I hypothesized that due to this, the protists that live attached to tubeworms bathed in fluid from active venting would be largely endemic, while protists living in the water above the vents would be largely cosmopolitan (also found elsewhere in the ocean). To test this, I collected samples from tubeworms and from the water above vents. I then extracted DNA from the samples and used polymerase chain reactions to select and make copies of a single conserved gene region, which was then sequenced. Two different sets of PCR primers were used, one that makes copies of all eukaryotic sequences and another that excludes animal (metazoan) sequences. It was also found that the universal primer set worked better than the metazoan-excluding primers to make copies of protist sequences from worm samples. Results of sequencing showed that more species were present in the water than on the tube worms, suggesting that more endemic species were present growing on the tube worms as hypothesized.

Introduction

Hydrothermal vent ecosystems are unique habitats subject to extremes in environmental conditions. Venting occurs primarily at mid ocean ridge spreading centers and over hot spots where the earth's crust is thin and heated rock from the mantle is close to the seafloor. In these areas of the seafloor, newly erupted rock lacking a sediment layer allows seawater to be drawn into cracks in the permeable basalt towards the heated magma below. As the water heats, it expands and dissolves elements from the basalt, leading to high pressure and high temperature fluid trapped within the rock. This mineral and metal laden fluid then rises to the surface of the seafloor and escapes at high velocity through cracks in the rock. This fluid may exit the seafloor at temperatures of over to 400°C with a pH as low as 1, devoid of oxygen, and laden with high concentrations of toxic metals (Kelley et al., 2002).

Carbon fixation in vent environments is powered by chemosynthesis carried out by autotrophic prokaryotes, which are in turn consumed by protists propagating the transfer of carbon up trophic levels. In the upper ocean, heterotrophic protists feed on phototrophic primary producers and heterotrophic prokaryotes. In the deep ocean, protists take part in a food chain that begins with chemical energy rather than sunlight. In the absence of sunlight, autotrophic organisms preyed on by protists must find other sources of chemical energy to utilize. Many of the chemicals present in hydrothermal fluid are unstable and in high energy molecular arrangements, making them ideal substrates for chemosynthetic organisms (bacteria and archaea) to use as electron donors or acceptors to fix carbon from carbon dioxide (Amend et al., 2011). Examples of reduced chemical species utilized as electron donors by specialized chemosynthetic organisms include hydrogen sulfide, Fe^{2+}, hydrogen gas and methane (Dick 2019).

Plumes of mixed vent fluid and seawater are enriched in total organic carbon (TOC) and prokaryotic cell counts relative to nearby background seawater, indicating significant carbon fixation in hydrothermal fluid and hydrothermally influenced seawater (Bennet et al., 2013). The productive ecosystems supported by chemosynthetic prokaryotes provide a high concentration of prey for heterotrophic protists to feed on (Hu et al., 2021). The same study conducted grazing incubations with hydrothermal fluids from five venting sites at Axial Seamount and found that protists consumed prokaryotic prey at 700 to 1,828 cells/mL/hour in diffuse vent fluids compared to 255 cells/mL/hour in background seawater. This study estimated that grazing by protists

accounts for 28 to 62% of the daily prokaryote biomass turnover daily, making heterotrophic protists an important source of labile carbon in the deep ocean and a potential link to higher trophic levels and overall carbon cycling at vents.

While hydrothermal fluid is rich in chemicals that support a productive ecosystem, it is also low in oxygen, high in temperature and low in pH; organisms that live in and around hydrothermal flow must be adapted to survive these selective conditions and vent communities may support a high proportion of endemic organisms (Sievert & Vetriani, 2012; Hu et al., 2022). An investigation into prokaryote concentrations at Guaymas Basin vents found cell concentrations in vent plumes (mixed vent fluid and seawater beginning ~1m from the vent orifice) to be around twice that of background nearby seawater (7 x 10^4 cell/mL and ~4 x 10^4 cell/mL respectively) (Hu et al., 2021). However, vent fluid itself was more variable, with concentrations of 1-5 x 10^4 cells/mL. This could indicate that vent plumes are more accessible for chemosynthetic organisms, as they have a lower temperature and higher oxygen content than high-temperature vent fluid, which may contain more adapted organisms.

Protistan communities in hydrothermal vent fluid and vent plumes have also been shown to differ from background seawater (Hu et al., 2022). Similarly, Sauvadet et al. in 2010 showed differences in community composition between deep seawater and vent fluids. The proportions of protists that are endemic to vent environments versus those that have a wider deep-sea distribution (cosmopolitan organisms) have been investigated in several studies. The total number of different species variants as defined by ASVs (Amplicon Sequence Variants) found to be “resident” in vent fluids was several-fold higher than in the cosmopolitan population (4,236 resident versus 535 cosmopolitan ASVs) (Hu et al. 2021). This was further supported by a 2022 study at the same sites showing 65% of ASVs in diffuse vent fluid were resident and only 17.2% were cosmopolitan (Hu et al. 2022). This study also found high variability within vents at the same vent field, with only 177 of 3586 ASVs found across multiple sampling years at the same vents. Conversely, a paper from Murdock & Juniper in 2019 found that 95% of the OTUs (Operational Taxonomic Units; an operational definition of taxonomic groupings defined by sequence homology) at three vent systems had cosmopolitan distribution across a broad range of vent and non-vent habitats. This study also found the lowest species richness (number of species) in the highest temperature fluids, indicating restriction of protistan diversity in high temperature vent fluids.

While restriction of diversity and higher proportions of endemic species have been demonstrated in communities of free-living organisms at vents, comparatively little research has compared the diversity of organisms living attached to surfaces in constant hydrothermal flow. A comparison of prokaryotic diversity within vent fluid, background seawater, and sediments bathed in continuous hydrothermal fluid flow revealed that diversity and richness were higher in the vent fluid than background seawater, but much lower in the continuously bathed sediments (Sheik et al., 2015). This indicates that organisms exposed to constant hydrothermal flow are more adapted than free-living microbes in order to survive the high temperatures and chemically reduced conditions. The first investigations into diversity within protists living attached to sediments in vent environments revealed diverse and deeply branching lineages with unexpected richness (Edgcomb, 2002). A later study on protists in hydrothermally influenced sediments characterized by microbial mats found 6,954 OTUs dominated by ciliates (Pasulka et al., 2019). This study also found that cold sediments outside of hydrothermal flow had higher richness in species, suggesting that hydrothermal exposure restricts diversity and selects for adapted organisms. In addition, differences in community composition were found at the centimeter scale. Within hydrothermal environments, microhabitats can form at the centimeter scale as thermal and chemical gradients create unique microscale conditions changing with distance from the vent orifice (Dick, 2019). This creates potential for small-scale ecological zonation, with organisms adapted to specific conditions present at different distances from hydrothermal flow. In order to maintain optimal conditions, organisms adapted to specific conditions may attach to surfaces to maintain their location along the thermochemical gradient. Organisms that are free-living will not be able to control their position along this gradient, making it potentially disadvantageous to become highly adapted to specific chemical, oxic, or temperature conditions. Therefore, organisms living attached to surfaces in hydrothermal flow may be more likely to be endemics not able to survive outside vent conditions than those living in the vent plume or background seawater.

Living substrates such as the chitinous exteriors of tube and palm worms are constantly exposed to hydrothermal fluid flow and host to a variety of epibiotic (living on the exterior of the host) microbes, presenting an opportunity to study attached communities of protists living in hydrothermal flow (Cary et al., 1987). Tube worms (*Ridgeia piscesae)* are specially adapted mouthless, gutless hydrothermal Vestimentiferans that rely on a symbiotic relationship with

chemosynthetic bacteria inside their trophosomes (highly adapted stomachs) for energy (Forget et al., 2015). *R. piscesae* are found in diffuse and active hydrothermal flow and depend on hydrothermal fluid to provide their endosymbionts with a means for carbon fixation (Perez & Juniper, 2018). So far, no investigations have been conducted into the protist communities living as epibiota on the exterior chitinous surfaces of tubeworms. Similar research into microbial epibiota and endobiota (living inside the host) of protists at hydrothermal and methane seep environments have been conducted and provide insight into methodology and potential findings of my proposed research. In 1997, Cary et al. took scrapings of the exterior of a vent sulfide worm (*Alvinella pompejana*) and were able to identify several bacterial taxa present as epibiota on the exterior of the worm. Sampling of nearby surfaces revealed that many of these bacteria were not unique to *A. pompejana*, indicating that the worms served as a surface for the bacteria to grow on rather than a source of food. Gill tissue samples from two species of bivalves inhabiting a methane seep were analyzed for protist colonization using a PCR primer set that excludes most metazoans. This study was able to identify a novel Alveolate within the gill cavity after extracting DNA from gill tissues (Noguchi et al., 2013; Bower et al., 2004). Using the same primer set, DNA was extracted and analyzed from fluid from within the pallial cavity of bivalves at a hydrothermal vent. This study identified a community of protists living within the clams (Sauvadet et al., 2010). In 2018, a parasitic protist was identified within the gut of a hydrothermal scale worm using microscopy and molecular techniques (Iritani et al., 2018). Surface scrapings of a squat lobster inhabiting hydrothermal vents was found to harbor a prokaryotic community distinct from that of nearby seawater, indicating that surface attached communities are different from those present in hydrothermal fluid and surrounding seawater (Leinberger et al., 2022). Considered together, these studies indicate that microbial communities living attached to larger metazoans in hydrothermal flow are different from that of nearby fluids but may not represent symbionts or parasites, rather utilizing the metazoan as an incidental surface on which to anchor. They also demonstrate the feasibility of isolating protists from metazoans using unique primers. I hypothesize that protistan communities living as epibiota on *Ridgeia piscesae* (tube worms) will have a lower community diversity than nearby vent plume samples because organisms attached to surfaces subjected to constant hydrothermal flow will need to be more highly adapted to the high temperatures, low pH and high concentrations of toxic chemicals to survive the selective and specific conditions of venting hydrothermal fluid

than organisms in the water column. In this study, I will compare the diversity, richness and community composition of protists living attached to the chitinous tubes of *Ridgeia piscesae* to free-living communities in nearby vent plumes at Axial Seamount using 18S rRNA genes to add to our understanding of community structure in vent ecosystems.

Methodology

Sample Collection

Samples were collected on the VISIONS research cruise aboard the *R/V Thomas G. Thompson* in August 2022. Targeted collection sites were hydrothermal vents in the ASHES Vent Field within the caldera of Axial Seamount, an undersea volcano along the spreading center of the Juan de Fuca plate off the coast of Oregon. The venting site Inferno, a large sulfide mound vent tower, was selected for worm sampling. Inferno vent is a high-temperature sulfide vent with an internal fluid temperature of ~300°C, a pH of ~3.5 and a high concentration of hydrogen sulfide (Butterfield et al.1990). Sample collection was conducted with the *ROV ROPOS* using the onboard Niskin bottle, ROV manipulators, and hydraulic BioBox. Samples were collected from two venting sites at ~1500m depth: one in low-velocity hydrothermal flow at the base of Inferno ("Inferno Diffuse") and one in high velocity flow on the flanks of Inferno ("Inferno Chimney"). Both sites had large assemblages of tube worms (*Ridgeia piscesae*) that were collected using the ROV manipulators and placed in separate BioBoxes. 4L Niskin bottles were deployed ~2m above each of the worm sites. An additional Niskin bottle was collected from ~2m above the topmost vent of Inferno ("plume" sample from "Inferno Chimney"). This third site was selected to gain an additional replicate of vent water with a high level of hydrothermal fluid input.

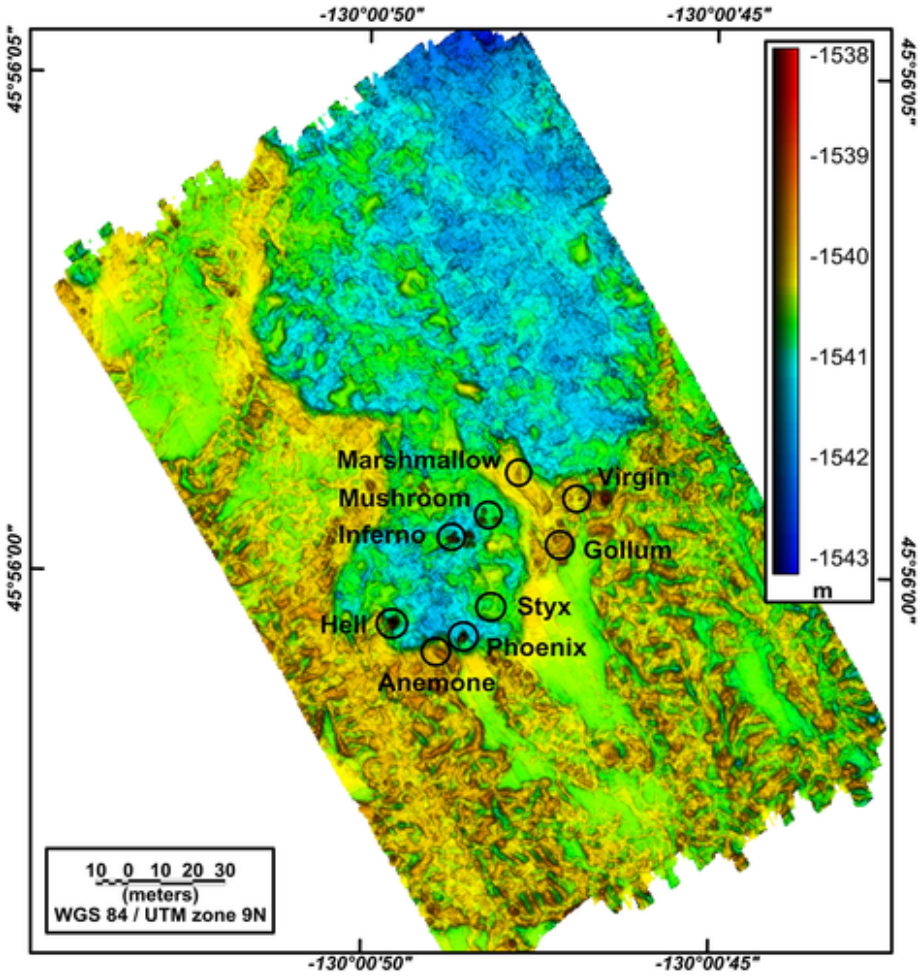


Figure 1. Bathymetry of ASHES vent field (0.25m resolution) with Inferno vent marked in the center of the figure. Adapted from Fig. 2 in Caratori et al., 2016.

Once brought onboard, worm samples were cut into sections and the soft body removed from the chitinous tube using a sterile scalpel to reduce worm tissue in the scrapings. Microbial growth was scraped from the outer surface of the tube sections using a sterile scalpel and the scrapings placed into a Falcon tube. Two tube scrapings were preserved from worms at the Inferno Chimney and two from worms at the Inferno Diffuse site (Table 1). All water samples from the Niskin bottles were filtered through sterile 0.8 micron mixed cellulose ester filters using a vacuum pump, with two replicates per Niskin bottle (Table 1). All samples were immediately placed into Ambion RNALater and stored at 4°C for 24hrs before being frozen and stored at -20°C. Materials and intermediate storage containers were cleaned with 10% HCl before use or purchased as sealed and sterile.

Sample Name	Sample Type	Site	Universal Primers	Non-Metazoan Primers
Hydrothermal Worm 1	Worm Scraping	Inferno Chimney	HW1U	HW1N
Hydrothermal Worm 2	Worm Scraping	Inferno Chimney	HW2U	HW2N
Hydrothermal Water 1	Niskin Water	Inferno Chimney	H1U	H1N
Hydrothermal Water 2	Niskin Water	Inferno Chimney	H2U	N/A
Diffuse Worm 1	Worm Scraping	Inferno Diffuse	DW1U	DW1N
Diffuse Worm 2	Worm Scraping	Inferno Diffuse	DW2U	N/A
Diffuse Water 1	Niskin Water	Inferno Diffuse	D1U	D1N
Diffuse Water 2	Niskin Water	Inferno Diffuse	D2U	N/A
Plume 1	Niskin Water	Inferno Chimney	P1U	P1N
Plume 2	Niskin Water	Inferno Chimney	P2U	N/A

Table 1. Sample names, abbreviated lab names for each sample with each primer set, sample types and sample collection site.

DNA Extraction and PCR

RNALater was removed from the samples using a vacuum pump onto a 0.4 micron PCTE filter. Filters were then placed into a 2mL Eppendorf DNALoBind tube for bead beating. DNA was extracted from the filters and tube worm scraping samples using a modified protocol for the Qiagen DNeasy Blood and Tissue extraction kit (Appendix 1). Sample DNA concentrations were quantified using a QuBit. The V4 and V5 hypervariable regions of the 18S rRNA gene were then amplified using PCR (Fig. 2) (Appendix 2). Universal eukaryotic primers Euk-F-566 (5'-CAGCAGCCGCGGTAATTCC-3') and Euk-R-1200 (5'-CCCGTGTTGAGT CAAATTAAGC-3') were used to amplify all samples (Hadziavdic et al., 2014). The primer set Euk581-F (5′-GTGCCAGCAGCCGCG-3′) and Euk1134-R (5′-TTTAAGTTTCAGCCTTGCG-3′) has been shown to be a "universal non-metazoan" primer set that excludes most metazoans to selectively amplify protists sequences in the same sample (Moon-van der Staay et al. 2000; Bower et al., 2004) . These primers were used to amplify DNA from the tube worm scraping samples to minimize amplification of worm DNA. Additionally,

PCR was carried out using the non-metazoan primer set on a DNA aliquot of the water samples and compared to the universal primers to validate the ability of non-metazoan primers to capture the full range of microbial eukaryote taxa present. Thermocycler protocols and master mix composition were adjusted from those described in Del Campo et al., 2019 to optimize amplification using the non-metazoan primer set (Appendix 2). Nextera adapter sequences were attached to the fragments (Fwd: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3', Rev: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3'); see Appendix 2 for thermocycler protocols. Bead cleaning was carried out using Omega Biotek Mag-Bind TotalPure NGS beads. Samples were then amplified again and indexed with Miseq Nextera adapters, then underwent a secondary bead cleaning to remove fragments below ~300bp. Products were pooled for each marker and all markers were pooled into a final library. The library was loaded onto the MiSeq at a concentration of 8pM with a spike in of 15% PhiX. PhiX is a viral gene library added to balance base composition and as an internal quality control. 2 × 300 bp paired end amplicon sequencing was then carried out to sequence sample DNA.

Sample	DNA conc.	Sample	DNA conc.
HW1	189µg/mL	H1	0.649µg/mL
HW2	130µg/mL	H2	0.398µg/mL
DW1	13.3µg/mL	D1	1.06µg/mL
DW2	7.1µg/mL	D2	0.410µg/mL
		P1	1.07µg/mL
		P2	0.887µg/mL

Table 2. DNA concentrations of worm scraping and water samples from all sites.

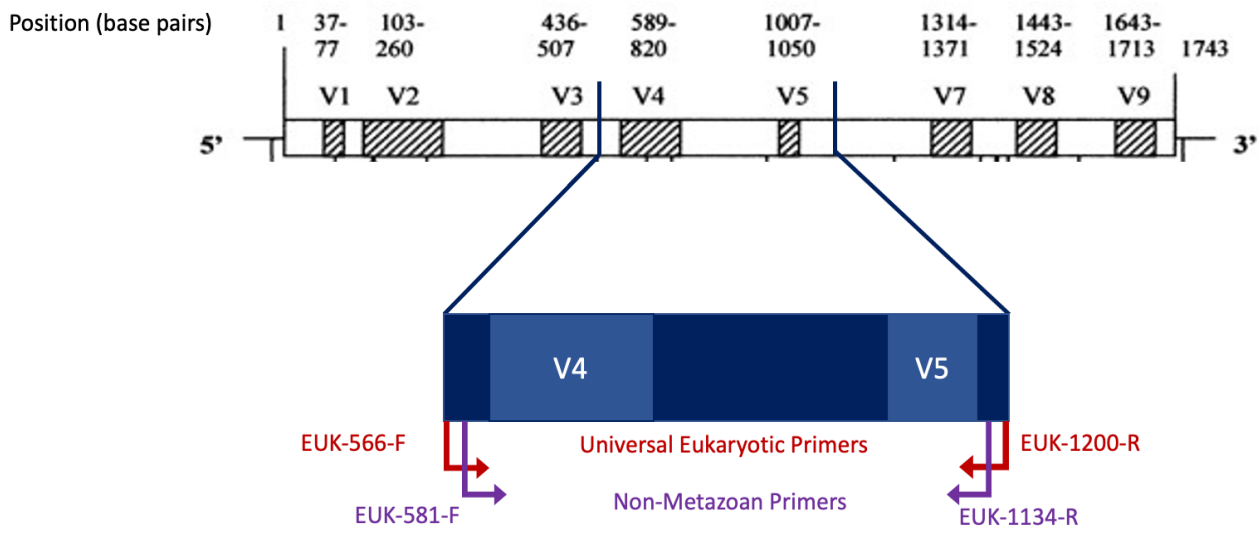


Figure 2. The 18S small ribosomal subunit gene with hypervariable regions highlighted. Expanded section including the V4 and V5 hypervariable regions highlights regions of amplification using universal and non-metazoan primers. Upper portion of gene image adapted from Machouart-Dubach et al., 2002.

Quality Control and Data Processing

Raw sequence files were demultiplexed using Miseq native software. Primer sequences were then removed using cutadapt version 4.4 on the command line (Martin, 2011). Cutadapt was also used to discard small fragments and reads that did not have complete primer sequences. Within RStudio, the package dada2 (version 3.26) was used for further processing up to taxonomic assignment (Callahan et al., 2016). Read quality profiles were plotted (Appendix III). Forward and reverse sequences were trimmed according to read quality profiles. Truncation length was chosen based on where in the read the quality score dropped below 30 (see Appendix III) (200bp forward and 120bp reverse for universal, 200bp forward and 140bp reverse for non-metazoan) then filtered to remove reads with high error rates (maxEE=2) in base transition. PhiX was removed from forward and reverse reads. Amplicon Sequence Variants (ASVs) were assigned using sample inference within dada2. Forward and reverse reads were joined using JustConcatenate=TRUE, as paired end reads did not have sufficient overlap to merge following truncation based on read quality profiles. Chimeras were removed (pooled method in dada2). Taxonomy was assigned within dada2 using the PR2 protist ribosomal database using bootstrap

values <80 (Guillou et. al, 2013) and graphical analysis was carried out in Rstudio and Microsoft Excel. Diversity was estimated using the Shannon-Weiner Diversity Index, calculated using $H = -\Sigma pi * \ln(pi)$. For this calculation, pi is the proportion of total protistan reads taken up by an given ASV. This was calculated by dividing the number of reads for a given protistan ASV at a site from a sample (worm or water) by the total number of protistan reads in that sample. Non-protist sequences were removed for graphing and analysis by excluding all metazoan and non-eukaryotic (bacteria and archaea) sequences.

Results

Primer Validation

The non-metazoan primer set did not amplify DNA successfully using the published starting protocol and adjustments were made to protocols to ensure complete amplifications. Using Phusion High Fidelity Taq polymerase along with the protocol published in del Campo et al., 2019 did not lead to successful amplification. Adjusting the initial denaturing temperature from 98°C to 94°C, the annealing temperature from 51.1°C to 60.0°C and the Taq from Phusion High Fidelity to Apex Taq all improved rates of successful amplification in worm scraping and water samples (Table 3). In addition, the high concentration of DNA (>100µg/mL) (Table 2) in the worm scraping samples HW1 and HW2 initially led to poor amplification (Table 3). Diluting these samples 1:20 improved amplification. Increasing the volume of low concentration samples (<5µg/mL) (Table 2) also improved the strength of amplification. When Nextera adapters were added to the existing primers, raising annealing temperature improved amplification and reduced the number of dimers present (Table 3). However, a high concentration of dimers were present in all Nextera adapter amplifications.

Primer Set	**Initial denaturing temperature**	**Annealing temperature**	**Taq**	**DNA amount (in 10µL reaction volume)**	**Amplification**
Non-metazoan	*98°C, 30 sec*	*51.1°C, 30 sec*	*Phusion HF*	*1µL worm, 2µL water*	*Very weak bands for DW1 and DW2, No for the rest*
Non-metazoan	*98°C, 30 sec*	*51.1°C, 30 sec*	*Phusion HF Taq*	*1µL HW1, HW2 diluted 1:20, 1µL DW1, DW2 diluted 1:10, 2µL water*	*Yes for DW1 and DW2, No for the rest*
Non-metazoan	*98°C, 30 sec*	*51.1°C, 30 sec*	*Phusion HF Taq*	*1µL HW1, HW2 diluted 1:20, 1µL DW1, DW2 diluted 1:10, 4µL water*	*Yes for DW1 and DW2, No for the rest*
Non-metazoan	*94°C, 2 min*	*60°C, 30 sec*	*Apex Taq*	*1µL HW1, HW2 diluted 1:20, 1µL DW1, DW2, 4µL water*	***Yes for all samples***
Non-metazoan with Nextera adapters	*94°C, 2 min*	*67°C, 30 sec*	*Apex Taq*	*1µL HW1, HW2 diluted 1:20, 1µL DW1, DW2, 4µL water*	*No, all dimers*
Non-metazoan with Nextera adapters	*94°C, 2 min*	*65°C, 30 sec*	*Apex Taq*	*1µL HW1, HW2 diluted 1:20, 1µL DW1, DW2, 4µL water*	***Yes, but with a lot of dimers***

Table 3. Trials carried out to determine optimal PCR protocol for amplifying with non-metazoan primers. "Water" refers to all water samples. See Appendix 2 for full thermocycler and master mix protocols.

Successful amplification using non-metazoan primers resulted in a band at around ~450 base pairs (Fig. 3). There were additional weaker bands for some worm scraping samples at >1000bp. Water samples had a secondary band around ~750bp. A secondary band ~250bp can be observed for two of the worm scraping samples.

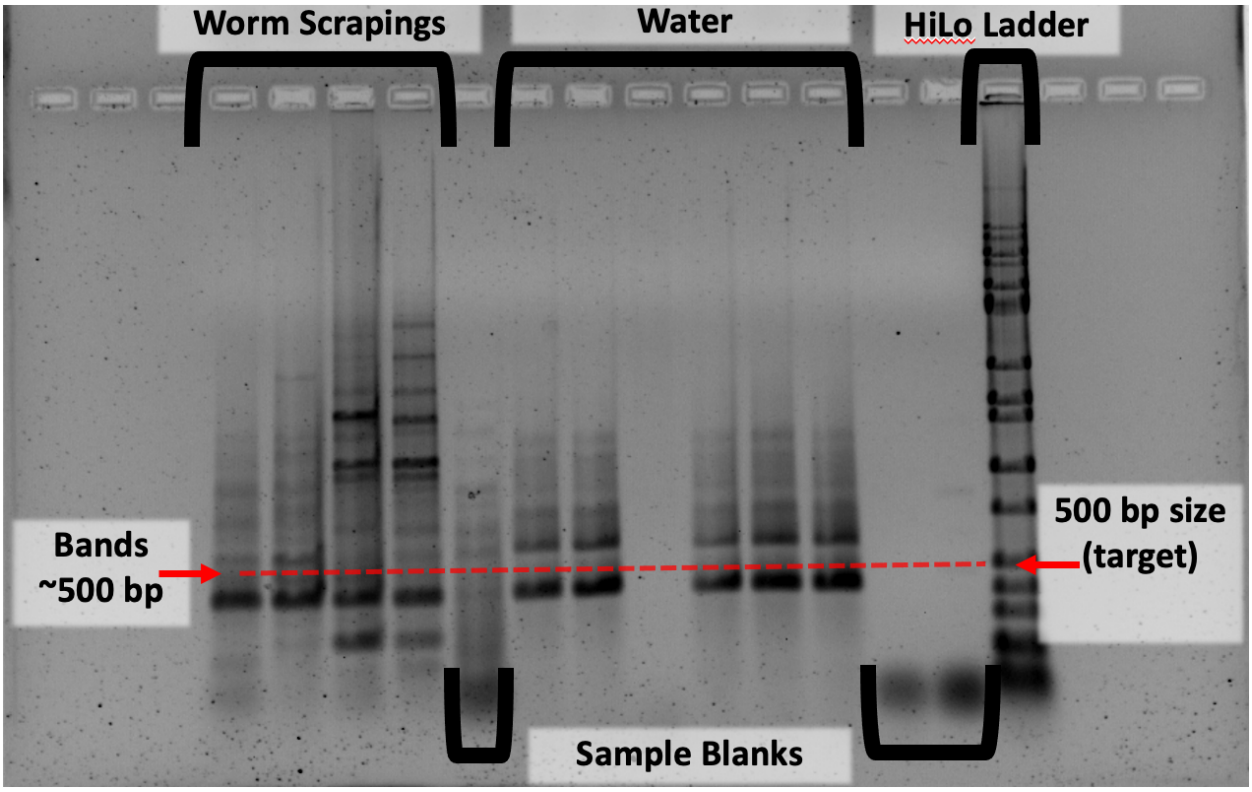


Figure 3. Bands of electrophoresis size sorted DNA, stained with SYBR, on 1% agarose gel under UV light. Visualizes size of DNA fragments resulting from PCR amplification of vent samples using Universal Non-Metazoan Primers (UNonMet) Euk-581-F and Euk-1134-R. Location of samples, blanks, and a size reference ladder in base pairs (bp) are shown.

ASV Richness

For protist sequences in samples amplified with non-metazoan primers, ASV richness (number of ASVs) was lower in the worm scraping sample than the water at both sites (Fig. 4a). For protist sequences in samples amplified with universal primers, ASV richness was lower in the worm scraping sample for all sites (Fig. 4b). Diversity of protists estimated using the Shannon-Weiner Diversity Index was higher for all water samples than worm samples (Fig. 5). This trend was consistent across sample sites and primer type. Diversity was similar between primer sets for the same sample, with the exception of Inferno Chimney (1.7 vs. 0.6 for universal and non-metazoan primers respectively). Diversity was consistent across water samples (+/-

0.55) regardless of site (Fig. 5a). For worm scraping samples, the range of diversity across site and samples types was broader (Fig. 5b).

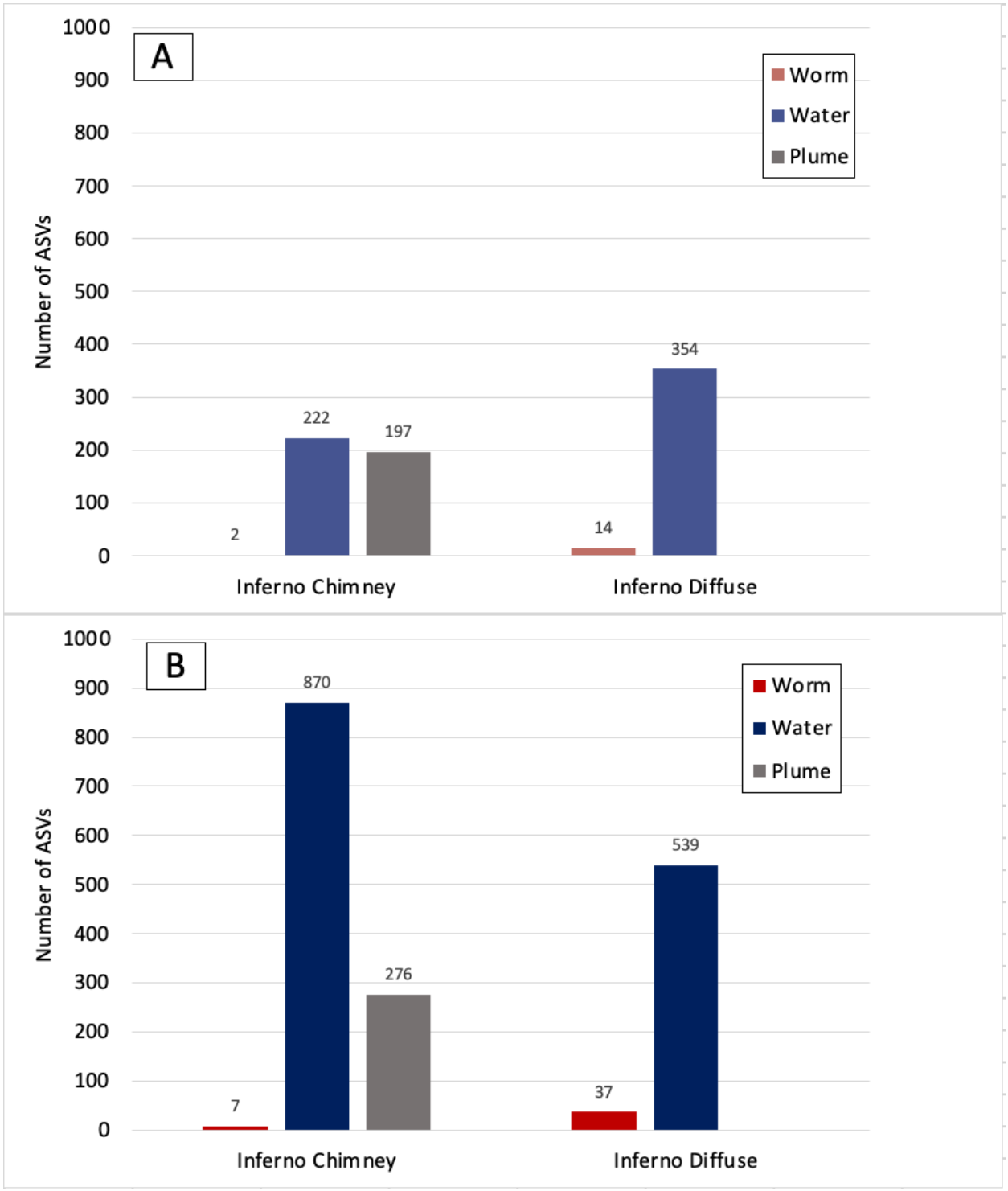


Figure 4. Amplicon Sequence Variant counts for worm scraping, water and plume samples at both sites. All non-protist sequences were removed prior to plotting. A: Samples amplified with

metazoan excluding primers. Samples HW2N, H1N, P1N, DW1N, and D1N plotted. B: Samples amplified with universal primers. Samples HW2N, H1N, P1N, DW1N, and D1N plotted.

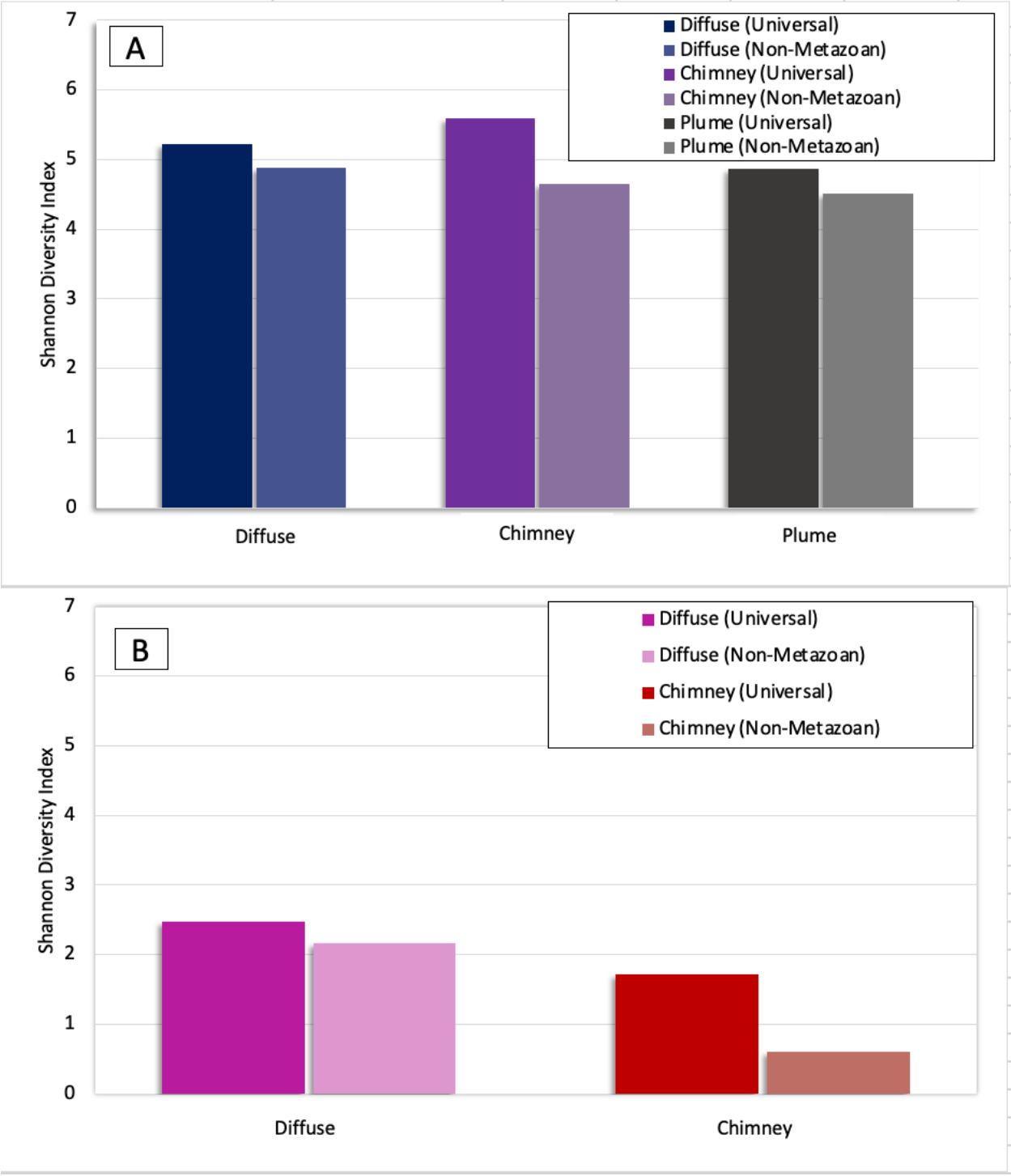
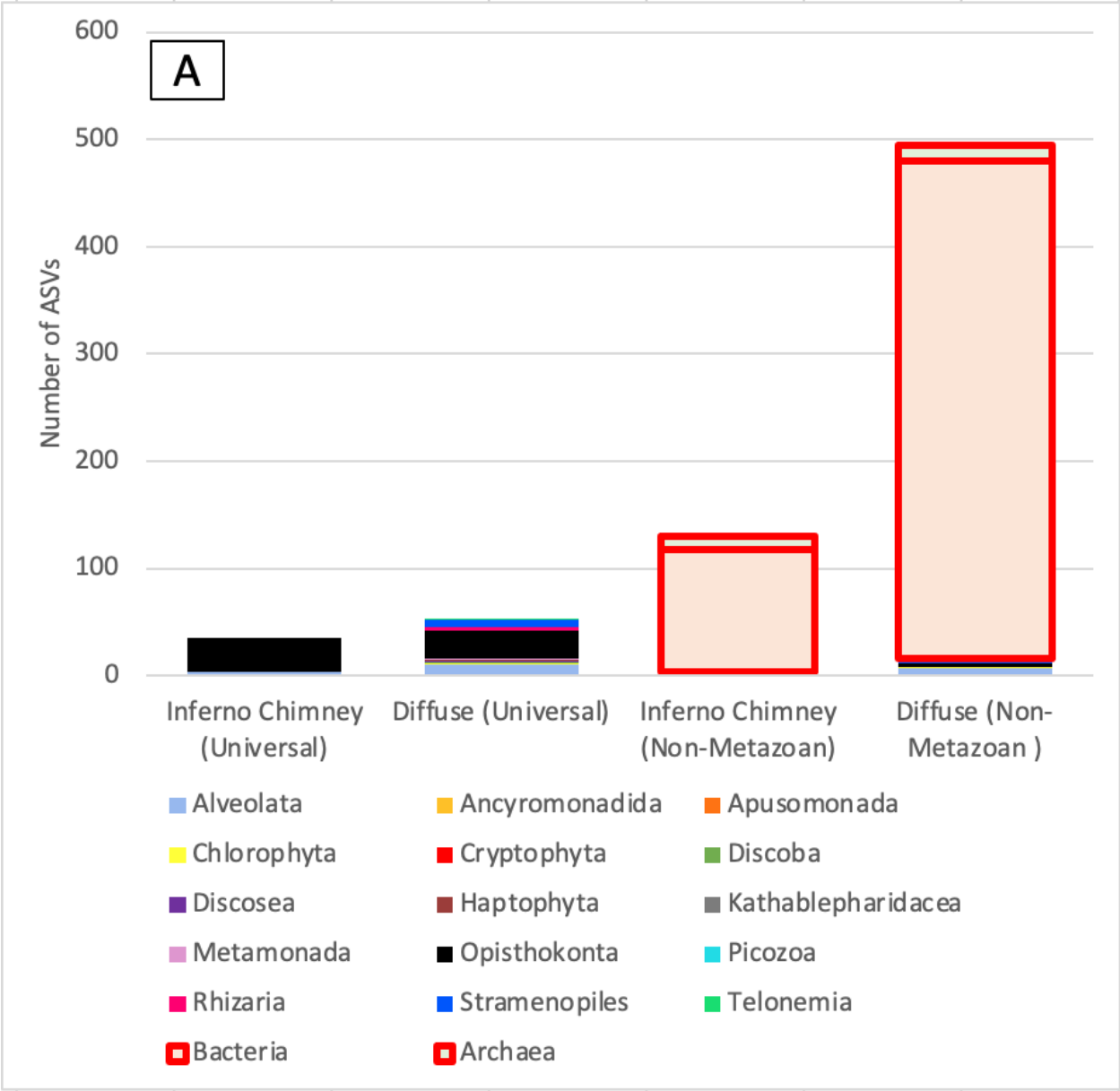


Figure 5. Comparison of Shannon Diversity Index for protist sequences amplified with universal vs. non-metazoan primers for water (A) and worm scraping (B) samples.

Taxonomic Conservation Between Primers

The distribution of taxa within the same samples was different between universal and non-metazoan primers, largely due to inclusion of prokaryotes with non-metazoan primers and inclusion of metazoans with universal primers. When non-metazoan primers were used on worm samples, a large number of bacterial and archaeal sequences were amplified in the worm scraping samples and relatively few protists by comparison, with 127/130 and 479/495 prokaryotic ASVs at chimney and diffuse sites respectively (Fig. 6a). Bacterial and archeal sequences represented a minority of ASVs in the water samples (117/339 for Chimney and 120/479 for Diffuse) (Fig. 6b). After prokaryotic ASVs were excluded, relative composition of protists was similar between universal and non-metazoan primers (Fig. 7b). Using universal primers, the supergroup Opisthokonta (which contains metazoans) was highly represented in worm scraping samples from both sites (Fig. 7a). A large number of ASVs from the chimney and diffuse worm scrapings were metazoan in origin (31/35 and 24/53 respectively), with *Riftia* making up 66% percent of reads in diffuse worm scrapings and 47% in chimney worm scrapings. Comparatively, the ASVs in the water samples contained few Opisthokonta with universal primers (71/935 for Chimney and 53/584 for Diffuse) (Fig. 7b). The dominant Supergroups in these samples were Alveolates (501/935 for Chimney and 314/584 for Diffuse), followed by Rhizaria (176/935 for Chimney and 108/584 for Diffuse).


A
600
500
400
300
200
100
0
Number of ASVs
Inferno Chimney (Universal)
Diffuse (Universal)
Inferno Chimney (Non-Metazoan)
Diffuse (Non-Metazoan)
Alveolata
Ancyromonadida
Apusomonada
Chlorophyta
Cryptophyta
Discoba
Discosea
Haptophyta
Kathablepharidacea
Metamonada
Opisthokonta
Picozoa
Rhizaria
Stramenopiles
Telonemia
Bacteria
Archaea

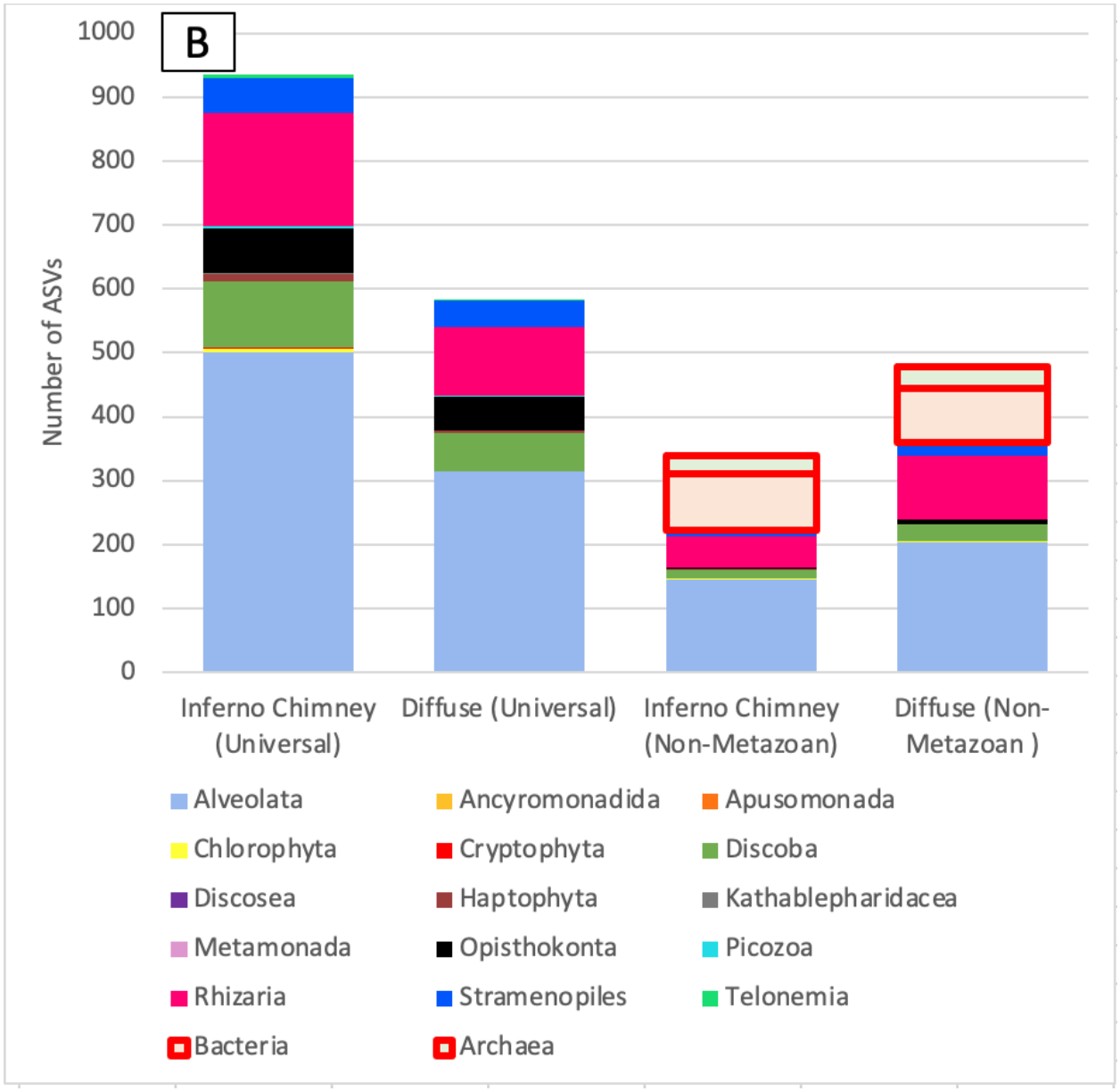
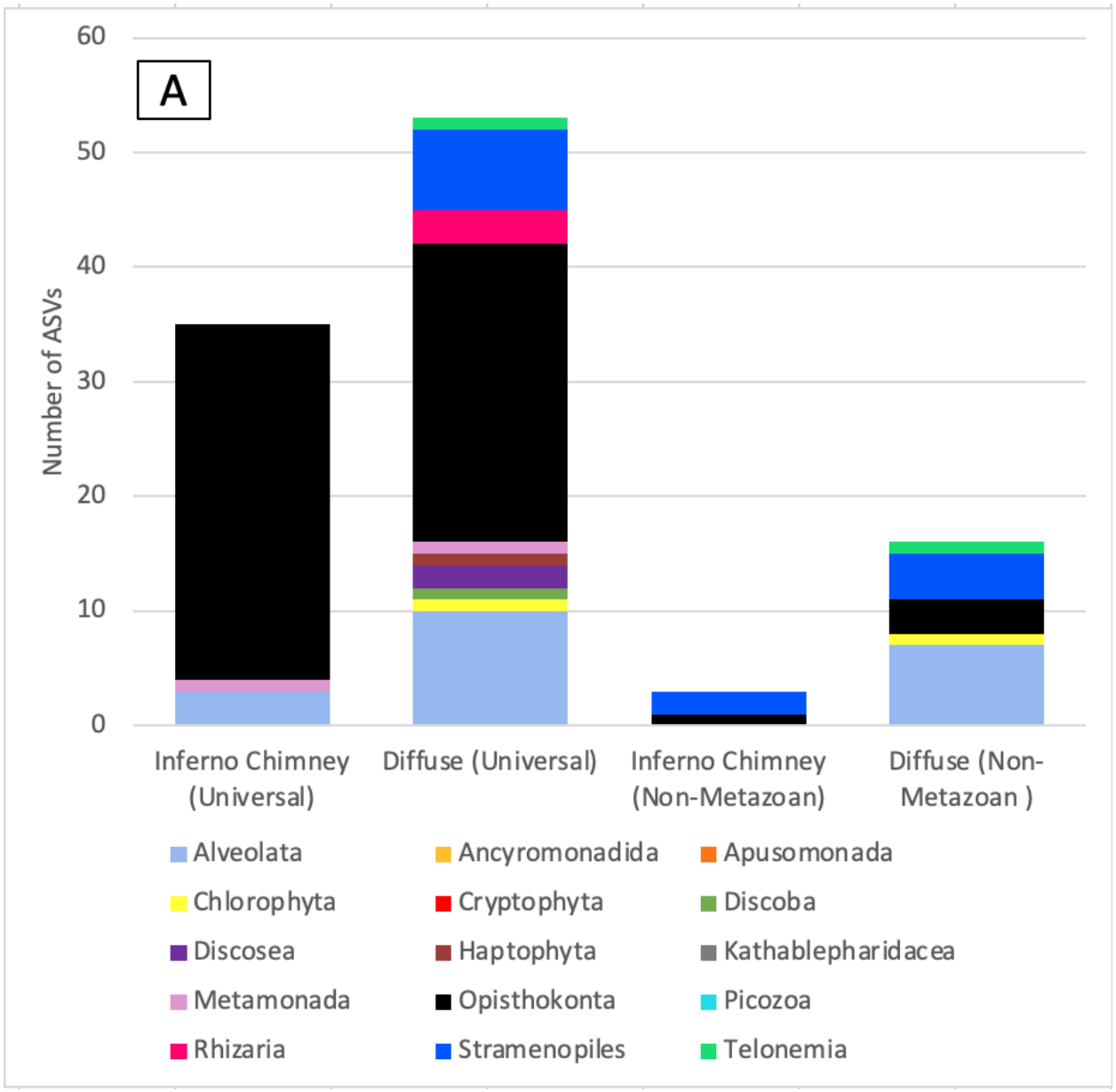


Figure 6. Comparison of Supergroups identified in worm scraping (A) and water (B) samples from all sites using universal vs. non-metazoan primers, prokaryotes included.


A
60
50
40
30
20
10
0
Number of ASVs
Inferno Chimney (Universal)
Diffuse (Universal)
Inferno Chimney (Non-Metazoan)
Diffuse (Non-Metazoan)
Alveolata
Ancyromonadida
Apusomonada
Chlorophyta
Cryptophyta
Discoba
Discosea
Haptophyta
Kathablepharidacea
Metamonada
Opisthokonta
Picozoa
Rhizaria
Stramenopiles
Telonemia

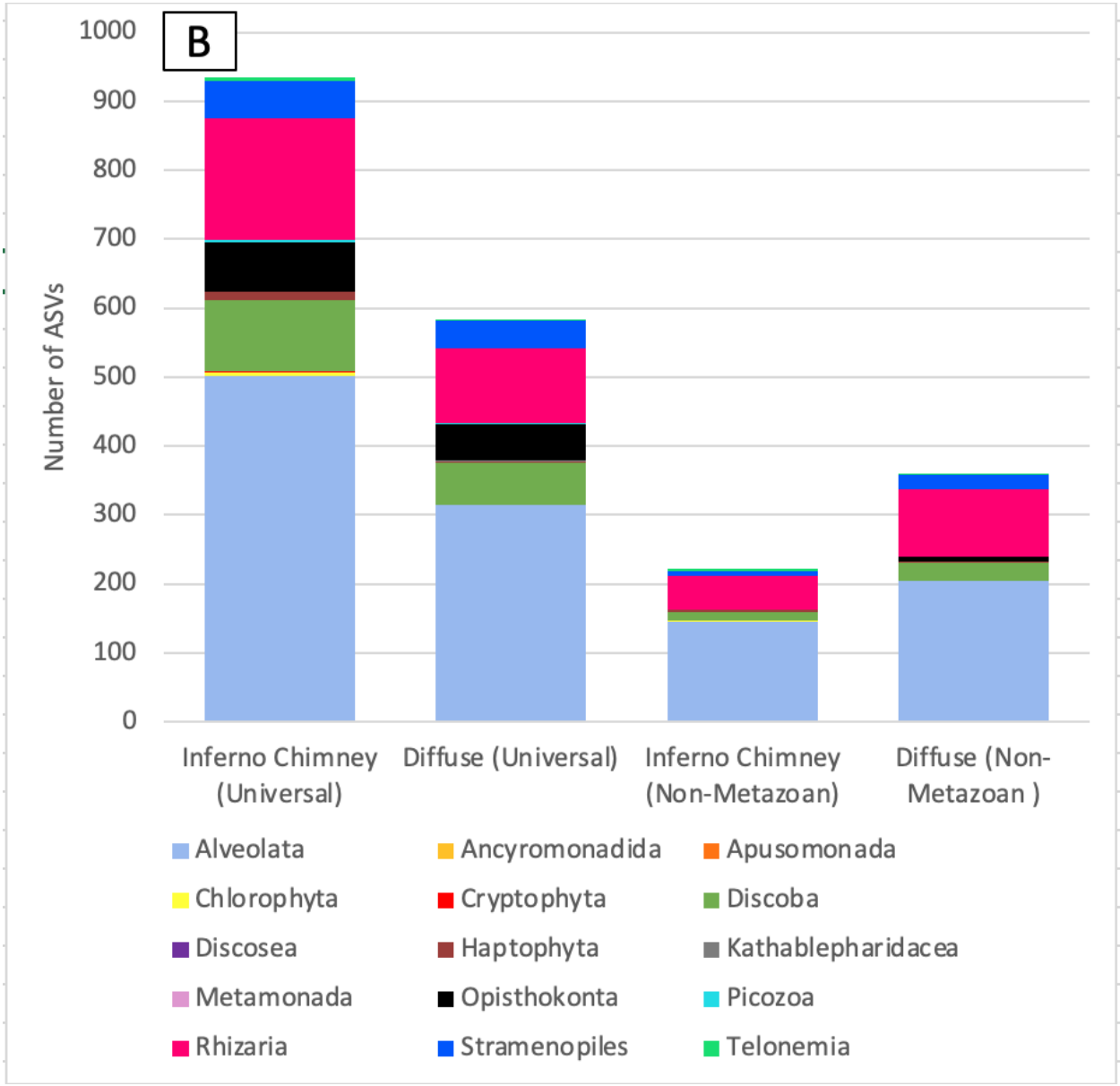


Figure 7. Comparison of Supergroups identified in worm scraping (A) and water (B) samples from all sites using universal vs. non-metazoan primers, prokaryotes excluded.

Discussion

Protistan ASV richness and diversity was lower in worm scraping samples than water samples, suggesting a more restricted community living on tube worms than in the water column above. The overall number of protistan ASVs at the tube worm microbial mats (worm scraping

samples) was lower than the number of ASVs in paired and plume water samples (Fig. 4), supporting the hypothesis that species richness would be lower growing on tubeworms. These results support the assumption endemism is likely higher at worm sites than in the water column, however further extrapolation cannot be confidently undertaken given the low number of replicates and sample sites and the absence of non-hydrothermally influenced deep sea water samples,. This reduction in protist ASVs from more hydrothermally influenced samples vs. water samples with lower hydrothermal influence was also found by Murdock et al. in 2019. A similar reduction in richness had also previously been observed for prokaryotes in sediments bathed in hydrothermal flow compared to nearby water samples (Sheik et al., 2015). In 2019, Pasulka et al. found that protists living in sediment with a greater degree of hydrothermal influence had a lower species richness than those living in less hydrothermally influenced sediments. This finding is replicated in my data, with fewer ASVs present attached to tube worms (ie. greater hydrothermal influence). Diversity in ASVs found after amplification with both sets of primers was higher for water samples than worm scraping samples (Fig. 5). Though true community diversity is difficult to reliably estimate, Shannon diversity is a stable method for estimating microbial diversity from rRNA genes (He et al., 2013) and the results of the diversity analysis further support the finding that fewer protists live attached to tube worms than in the water column. Given that no previous research has investigated the presence of protists living attached to worms or other living substrates in hydrothermal flow, the finding that protist sequences can be isolated from worm scraping samples represents a discovery in and of itself. Edgcomb et al. in 2002 found protists living in hydrothermally influenced sediments, the first investigation into their presence on hard substrates (as opposed to free living in the water column). Further study is recommended to investigate the relationship between attachment to surfaces in hydrothermal flow and endemism. Some of these ASVs likely represent new taxa as evidenced by the lack of Family and lower levels of classification for many ASVs, and confirmation of living protists attached to tubeworms using RNA sequencing (as RNA is only transcribed by active organisms rather than DNA present in detritus) could further strengthen the conclusion that diverse protistan assemblages live and grow in direct hydrothermal flow.

Non-metazoan primers were not shown to be effective in isolating protist sequences from worm scraping samples; while metazoans were largely excluded, many prokaryotic sequences were amplified and number of protistan ASVs was reduced compared to universal eukaryotic

primers. The relationship between overall (protistan and non-protistan) number of ASVs and sample type was dependent on the primer set used to amplify sample DNA, with non-metazoan primers returning much higher numbers of ASVs for worm scraping samples before prokaryotic sequences were removed (Fig. 6). Additional investigation into the taxonomic composition of the returned ASVs for each primer set offered possible explanations for the variability within ASV richness results. Using universal primers, up to 66% of the reads in worm scraping samples were from *Riftia*, the species of tube worm scraped for this study (Fig. 7). In spite of this high proportion of worm DNA, the number of protistan ASVs in the water samples was between 276-870 compared to the non-metazoan primer amplified sequences with 197-354 ASVs with 7-37 and 1-14 ASVs for worm samples (Fig. 4). This suggests that universal primers are better at capturing the true diversity of protistan ASVs in hydrothermal samples despite capturing metazoan ASVs. The non-metazoan primer set was shown to amplify a large number of prokaryotic sequences (Fig. 5), counter to its intended purpose of amplifying non-metazoan protist sequences. In the paper first describing the non-metazoan primers, bacterial sequences around 400bp (close to the target length) were amplified by the primer set using annealing temperatures below 49°C (Bower et al., 2004). This should have been avoided for samples in this study by using an annealing temperature of 65°C for samples with non-metazoan + Nextera adapters. In general, higher annealing temperatures increase primer specificity and reduce bias (Obradovic et al., 2013). Bower et al. also described a weak band around 600bp and 1200bp when bacterial DNA was amplified using non-metazoan primers. It is more likely that these larger bands, which can be seen in the worm scraping and water samples (Fig. 6), represented prokaryotic amplification. There were comparatively fewer protistan ASVs in the non-metazoan primer amplified sequences compared to the universal primers for worm scraping and water samples (Fig. 4). Given this, non-metazoan primers may not be a viable solution for excluding metazoan sequences from worm scraping samples in the future. One possible solution for removing prokaryote sequences from non-metazoan primer samples would be to implement a nested PCR. Initial amplification could be undertaken with universal eukaryotic primers (which did not amplify any prokaryotic sequences) Euk-F-566 and Euk-R-1134, then a secondary PCR using non-metazoan primers Euk-F-581 and Euk-R-1134. Since the range of non-metazoan primers is a smaller portion of the 18S gene that fits within the region amplified by universal eukaryotic primers, secondary amplification should be possible. Without additional testing, it

cannot be determined whether the lower number of protist ASVs present in the non-metazoan primer amplified samples was due to lack of specificity or competition with a high number of prokaryotic sequences. Given that read depth was higher for worm scrapings amplified using non-metazoan primers (Appendix 2), it may be worthwhile to investigate using nested PCR as the product of non-metazoan primers is a shorter fragment that is more easily read and merged using NGS Miseq. Alternatively, since the universal eukaryotic primers did amplify a large number of protist sequences along with metazoans, it may not be useful to employ metazoan excluding primers for worm scraping samples in the future.

The non-metazoan primer set used in this study was designed to amplify the V4 hypervariable region of the 18S gene, however its use has not been widely adopted since its conception in 2004 leading to inconsistencies in published protocols (Bower, 2004; del Campo et al., 2019). In 2019, del Campo et al. used an annealing temperature of 51.1°C and Phusion HF Taq polymerase to amplify samples of coral and their associated microbiota using non-metazoan primers. Given that this was one of only four published papers utilizing the non-metazoan primer set and the most recent, I selected these conditions as a starting point for their amplification protocol. Initial runs using the above specifications failed to amplify significant amounts of DNA. The weak bands that were observed in the initial amplification were for the samples DW1N and DW2N, worm scraping samples from the Inferno diffuse site (Table 3). The first protocol modification carried out by I was to increase the concentration of DNA for low concentration samples, which did not alone affect the strength of amplification (Table 3). Next, the brand of Taq was shifted from Phusion HF to Apex, the same taq used for the universal primer. Thermocycler protocol was also adjusted to match that of the universal primer set, and a further increase in DNA concentration for the low concentration samples was utilized. This combination did result in successful amplification. Given these results, I recommend adjusting the PCR protocol to match that described above.

Conclusions

The most significant result is the discovery that protistan assemblages live attached to tube worms growing in direct hydrothermal flow. No prior studies have investigated protistan growth attached to living substrates at hydrothermal vents. This result can be used as a launching point for further investigations into protistan assemblages and diversity at hydrothermal vents. A secondary result that a higher diversity and number of ASVs are present in the water column above direct hydrothermal flow than growing attached to tube worms suggests a high proportion of endemic species specially adapted to live in hydrothermal flow are present on tube worms. The non-metazoan primer set was shown to amplify a large number of prokaryotic sequences when utilized with an adapted protocol but the universal eukaryotic primers were able to amplify protist sequences even on worm samples with a higher number of ASVs captured. Further research into endemism of taxa present on tube worms is recommended, as is the utilization of universal eukaryotic primers.

Acknowledgements

I would like to thank everyone involved in the collection and processing of water and worm samples including Andrew Paley, Mariela White and Anna Sulc, and the captains and crews of the *R/V Thompson* as well as the pilots and engineers for *ROV ROPOS*. I would also like to thank the Armbrust Lab, especially Megan Schatz and Rhonda Morales, for their invaluable mentoring, lab resources and assistance with sample processing. I would like to thank Dr. Ginger Armbrust and Dr. François Ribalet for their guidance, feedback and mentorship. Additionally, I owe a great deal to Dr. Megan Shaffer and Alexandria Im of the Kelley Lab for indexing, library creation, sequencing and bioinformatics assistance. This research was funded by the University of Washington senior thesis budget (65-8934) and the NSF VISIONS project.

References

Amend, J. P., McCollom, T. M., Hentscher, M., & Bach, W. (2011). Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted in different rock types. *Geochimica et Cosmochimica Acta*, *75*(19), 5736–5748. https://doi.org/10.1016/j.gca.2011.07.041

Bower, S. M., Carnegie, R. B., Goh, B., Jones, S. R. M., Lowe, G. J., & Mak, M. W. S. (2004). Preferential PCR Amplification of Parasitic Protistan Small Subunit rDNA from Metazoan Tissues. *Journal of Eukaryotic Microbiology*, *51*(3), 325–332. https://doi.org/10.1111/j.1550-7408.2004.tb00574.x

Butterfield, D. A., Massoth, G. J., McDuff, R. E., Lupton, J. E., & Lilley, M. D. (1990). Geochemistry of hydrothermal fluids from Axial Seamount hydrothermal emissions study vent field, Juan de Fuca Ridge: Subseafloor boiling and subsequent fluid-rock interaction. *Journal of Geophysical Research: Solid Earth*, *95*(B8), 12895–12921. https://doi.org/10.1029/JB095iB08p12895

Caratori Tontini, F., Crone, T. J., de Ronde, C. E. J., Fornari, D. J., Kinsey, J. C., Mittelstaedt, E., & Tivey, M. (2016). Crustal magnetization and the subseafloor structure of the ASHES vent field, Axial Seamount, Juan de Fuca Ridge: Implications for the investigation of hydrothermal sites. *Geophysical Research Letters*, *43*(12), 6205–6211. https://doi.org/10.1002/2016GL069430

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Cary, S. C., Cottrell, M. T., Stein, J. L., Camacho, F., & Desbruyeres, D. (1997). Molecular Identification and Localization of Filamentous Symbiotic Bacteria Associated with the Hydrothermal Vent Annelid Alvinella pompejana. *Applied and Environmental Microbiology*, *63*(3), 1124–1130. https://doi.org/10.1128/aem.63.3.1124-1130.1997

Dick, G. J. (2019). The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. *Nature Reviews Microbiology*, *17*(5), 271–283. https://doi.org/10.1038/s41579-019-0160-2

del Campo, J., Pons, M. J., Herranz, M., Wakeman, K. C., del Valle, J., Vermeij, M. J. A., et al. (2019). Validation of a universal set of primers to study animal-associated microeukaryotic communities. *Environmental Microbiology*, *21*(10), 3855–3861. https://doi.org/10.1111/1462-2920.14733

Edgcomb, V. P., Kysela, D. T., Teske, A., de Vera Gomez, A., & Sogin, M. L. (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proceedings of the National Academy of Sciences*, *99*(11), 7658–7662. https://doi.org/10.1073/pnas.062186399

Forget, N. L., Perez, M., & Juniper, S. K. (2015). Molecular study of bacterial diversity within the trophosome of the vestimentiferan tubeworm Ridgeia piscesae. *Marine Ecology*, *36*(S1), 35–44. https://doi.org/10.1111/maec.12169

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H. C. F., Lara, E., Le Bescot, N., Logares, R. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small

Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, *41*(D1), D597–D604. https://doi.org/10.1093/nar/gks1160

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. *PLoS ONE*, *9*(2), e87624–e87624. https://doi.org/10.1371/journal.pone.0087624

He, Y., Zhou, B.-J., Deng, G.-H., Jiang, X.-T., Zhang, H., & Zhou, H.-W. (2013). Comparison of microbial diversity determined with the same variable tag sequence extracted from two different PCR amplicons. *BMC Microbiology*, *13*(1), 208. https://doi.org/10.1186/1471-2180-13-208

Hu, S. K., Herrera, E. L., Smith, A. R., Pachiadaki, M. G., Edgcomb, V. P., Sylva, S. P., Chan, E. W., Seewald, J. S., German, C. R., & Huber, J. A. (2021). Protistan grazing impacts microbial communities and carbon cycling at deep-sea hydrothermal vents. *Proceedings of the National Academy of Sciences*, *118*(29), e2102674118. https://doi.org/10.1073/pnas.2102674118

Hu, S. K., Smith, A. R., Anderson, R. E., Sylva, S. P., Setzer, M., Steadmon, M., Frank, K. L., Chan, E. W., Lim, D. S. S., German, C. R., Breier, J. A., Lang, S. Q., Butterfield, D. A., Fortunato, C. S., Seewald, J. S., & Huber, J. A. Hu, S. K., Smith, A. R., Anderson, R. E., Sylva, S. P., Setzer, M., Steadmon, M., et al. (2022.). Globally-distributed microbial eukaryotes exhibit endemism at deep-sea hydrothermal vents. *Molecular Ecology*. https://doi.org/10.1111/mec.16745

Iritani, D., Horiguchi, T., & Wakeman, K. C. (2018). Molecular Phylogenetic Positions and Ultrastructure of Marine Gregarines (Apicomplexa) Cuspisella ishikariensis n. gen., n.

sp. and Loxomorpha cf. harmothoe from Western Pacific scaleworms (Polynoidae). *Journal of Eukaryotic Microbiology*, *65*(5), 637–647. https://doi.org/10.1111/jeu.12509

Kelley, D. S., Baross, J. A., & Delaney, J. R. (2002). Volcanoes, Fluids, and Life at Mid-Ocean Ridge Spreading Centers. *Annual Review of Earth and Planetary Sciences*, *30*(1), 385–491. https://doi.org/10.1146/annurev.earth.30.091201.141331

Leinberger, J., Milke, F., Christodoulou, M., Poehlein, A., Caraveo-Patiño, J., Teske, A., & Brinkhoff, T. (2022). Microbial epibiotic community of the deep-sea galatheid squat lobster Munidopsis alvisca. *Scientific Reports*, *12*(1), 2675. https://doi.org/10.1038/s41598-022-06666-x

Machouart-Dubach, M., Lacroix, C., Vaury, C., Feuilhade de Chauvin, M., Bellanné, C., Derouin, F., & Lorenzo, F. (2002). Nucleotide structure of the Scytalidium hyalinum and Scytalidium dimidiatum 18S subunit ribosomal RNA gene: evidence for the insertion of a group IE intron in the rDNA gene of S. dimidiatum. *FEMS Microbiology Letters*, *208*(2), 187–196. https://doi.org/10.1016/S0378-1097(02)00445-7

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Murdock, S. A., & Juniper, S. K. (2019). Hydrothermal vent protistan distribution along the Mariana arc suggests vent endemics may be rare and novel. *Environmental Microbiology*, *21*(10), 3796–3815. https://doi.org/10.1111/1462-2920.14729

Noguchi, F., Kawato, M., Yoshida, T., Fujiwara, Y., Fujikura, K., & Takishita, K. (2013). A Novel Alveolate in Bivalves with Chemosynthetic Bacteria Inhabiting Deep-Sea Methane Seeps. *Journal of Eukaryotic Microbiology*, *60*(2), 158–165. https://doi.org/10.1111/jeu.12017

Obradovic, J., Jurisic, V., Tosic, N., Mrdjanovic, J., Perin, B., Pavlovic, S., & Djordjevic, N. (2013). Optimization of PCR Conditions for Amplification of GC‐Rich EGFR Promoter Sequence. *Journal of Clinical Laboratory Analysis*, *27*(6), 487–493. https://doi.org/10.1002/jcla.21632

Pasulka, A., S. K. Hu, P. D. Countway, K. J. Coyne, S. C. Cary, K. B. Heidelberg, and D. A. Caron. 2019. SSU-rRNA Gene Sequencing Survey of Benthic Microbial Eukaryotes from Guaymas Basin Hydrothermal Vent. Journal of Eukaryotic Microbiology **66**: 637–653. doi:10.1111/jeu.12711

Perez, M., and S. K. Juniper. 2018. Is the trophosome of Ridgeia piscesae monoclonal? Symbiosis **74**: 55–65. doi:10.1007/s13199-017-0490-7

Sauvadet, A.-L., Gobet, A., & Guillou, L. (2010). Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. *Environmental Microbiology*, *12*(11), 2946–2964. https://doi.org/10.1111/j.1462-2920.2010.02272.x

Sheik, C. S., Anantharaman, K., Breier, J. A., Sylvan, J. B., Edwards, K. J., & Dick, G. J. (2015). Spatially resolved sampling reveals dynamic microbial communities in rising hydrothermal plumes across a back-arc basin. *The ISME Journal*, *9*(6), 1434–1445. https://doi.org/10.1038/ismej.2014.228

Sievert, S., & Vetriani, C. (2012). Chemoautotrophy at Deep-Sea Vents: Past, Present, and Future. *Oceanography*, *25*(1), 218–233. https://doi.org/10.5670/oceanog.2012.21

Appendices

Appendix I: DNA Extraction Protocol (Modified Qiagen Blood and Tissue Extraction Kit Protocol)

- For worm scrapings, begin with a 15mL Falcon tube containing the tube scrapings and 15mL RNALater stored at -20°C. Samples were thawed for an hour before processing.
    - Pour RNALater containing tube scrapings into a vacuum rig over a 0.4µm PCTE filter and suction off the liquid
    - Into a sterile 50mL Falcon tube and add 15mL PBS to the empty 15mL tube and invert
    - Add the 15mL PBS to the 50mL tube with the RNALater to dilute then invert several times (RNALater is quite viscous and requires dilution prior to filtering).
    - Pour RNALater and PBS mixture into a vacuum rig over a 0.4µm PCTE filter and suction off the liquid
    - Rinse the vacuum rig with PBS to wash any particulates onto the filter
    - Vacuum a second time
    - Transfer 0.4µm filter to a sterile petri dish on ice, cut the filter into small strips and transfer them into a 2mL centrifuge tube containing glass beads
    - Add 720µL buffer ATL from the Qiagen Blood and Tissue kit
- For water samples, begin with a 15mL Falcon tube containing one 47mm mixed cellulose ester 0.8µm filter and 15mL RNALater stored at -20°C. Samples were thawed for an hour before processing.
    - Remove filter from RNALater using sterile forceps
    - Over a sterile petri dish on ice, cut the filter into small strips and transfer them into a 2mL centrifuge tube containing glass beads
    - Pour remaining RNALater into a sterile 50mL Falcon tube and add 15mL PBS to the empty 15mL tube and invert
    - Add the 15mL PBS to the 50mL tube with the RNALater to dilute then invert several times (RNALater is quite viscous and requires dilution prior to filtering).
    - Pour RNALater and PBS mixture into a vacuum rig over a 0.4µm PCTE filter and suction off the liquid

    - Transfer 0.4µm filter to the same petri dish and cut as with the original filter, add cut strips to the same 2mL tube as the original filter
    - Add 720µL buffer ATL from the Qiagen Blood and Tissue kit
- At this point, treatment of both sample types is identical and combined for the remaining steps.
    - Samples were bead beat for 50 seconds and transferred to a 56°C shaking incubator at 200rpm for 30 minutes
    - Bead beating and incubation were repeated
    - 80µL of proteinase K was added
    - Samples were placed back in the shaking incubator overnight at 56°C
    - The next day, samples were removed from the incubator and 650µL of sample was pipetted off (avoiding the glass beads) then transferred to a 1.5mL sterile centrifuge tube. Note: the viscosity of RNALater causes pipetting to be difficult, you may need to use a smaller pipette tip or spin down the tubes to extract the liquid
    - 1.5mL tubes were spun down for 1 minute at 13,000rpm
    - 600µL supernatant was transferred to a clean 2mL tube
    - 600µL buffer AL was added
    - Samples were vortexed for 5 seconds then incubated at 56°C for 10 minutes
    - 600µL ethanol was added
    - Samples were vortexed for 5 seconds
    - 500µL of sample was pipetted into a Qiagen spin column over a 2mL tube
    - Centrifuged at 8000rpm for 1 min, flow through discarded
    - Above 2 steps were repeated until all the sample was run through the spin column
    - Spin column was transferred into a new 2mL tube
    - 500µL buffer AW1 was added
    - Centrifuged at 8000rpm for 1 min, discard flow through
    - Repeat above 2 steps
    - Spin columns placed in new 2mL tubes
    - 500µL buffer AW2 added
    - Centrifuged at 14,000rpm for 3 minutes, discard flow through

- Repeat above 2 steps
- Spin column was placed into a clean 1.5 mL centrifuge tube
- 25µL RNase and DNase free water was added the center of the column
- Incubated at room temperature for 1 minute
- Spun down at 8000rpm for 1 minute, keep flow through
- Repeat above 3 steps for a total of 50µL water eluted
- Finish - DNA is in the water
- Quantify DNA

Appendix II: PCR Protocols

For universal primers (no adapters)

Master Mix (per sample)

MgCl2 50 mM	0.5µL
10x buffer	1µL
Forward primer 10mM	0.4µL
Reverse primer 10mM	0.4µL
dNTPs 8mM	0.5µL
Apex Taq DNA polymerase	0.075µL
Water	6.125µL
DNA (sample)	1µL 20x diluted HW1, HW2; 1µL 10x dilute DW1, DW2; 1µL H1, H2, D1, D2, P1, P2

Thermocycler Protocol

Temperature	**Time**	**Number of Cycles**
94°C	2 minutes	1
94°C	30 seconds	35
60°C	30 seconds	35
72°C	1 minute	35

72°C	10 minutes	1

For universal primers (with Nextera adapters)

Master Mix (per sample)

MgCl2 50 mM	1.25µL
10x buffer	2.5µL
Forward primer 10mM	1µL
Reverse primer 10mM	1µL
dNTPs 8mM	1.25µL
Apex Taq DNA polymerase	0.1875µL
Water	12.8125µL
DNA (sample)	2.5µL 20x diluted HW1 +2.5µL H2O, HW2, 2.5µL DW1, DW2 + 2.5 µL H2O; 5µL H1, H2, D1, D2, P1, P2

Thermocycler Protocol

Temperature	**Time**	**Number of Cycles**
94°C	2 minutes	1
94°C	30 seconds	35
67°C	30 seconds	35
72°C	1 minute	35
72°C	10 minutes	1

For non-metazoan primers (no adapters)

Master Mix (per sample)

MgCl2 50 mM	0.5µL
10x buffer	1µL
Forward primer 10mM	0.4µL

Reverse primer 10mM	0.4µL
dNTPs 8mM	0.5µL
Apex Taq DNA polymerase	0.075µL
Water	3.125µL
DNA (sample)	1µL 20x diluted HW1, HW2 + 3µL H2O; 1µL 10x dilute DW1, DW2 + 3µL H2O; 4µL H1, H2, D1, D2, P1, P2

Thermocycler Protocol

Temperature	**Time**	**Number of Cycles**
94°C	2 minutes	1
94°C	30 seconds	35
60°C	30 seconds	35
72°C	1 minute	35
72°C	10 minutes	1

For universal primers (with Nextera adapters)

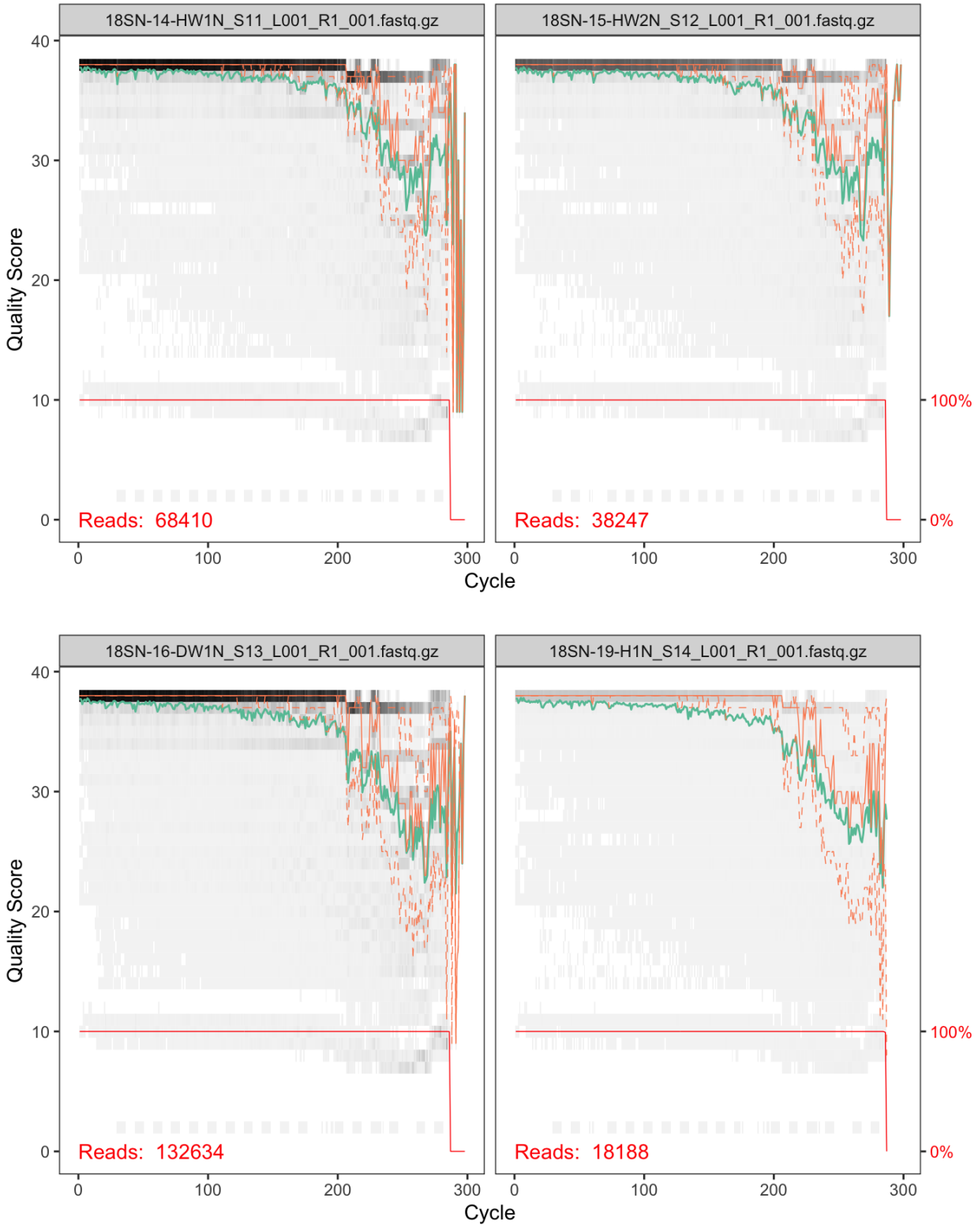
Master Mix (per sample)

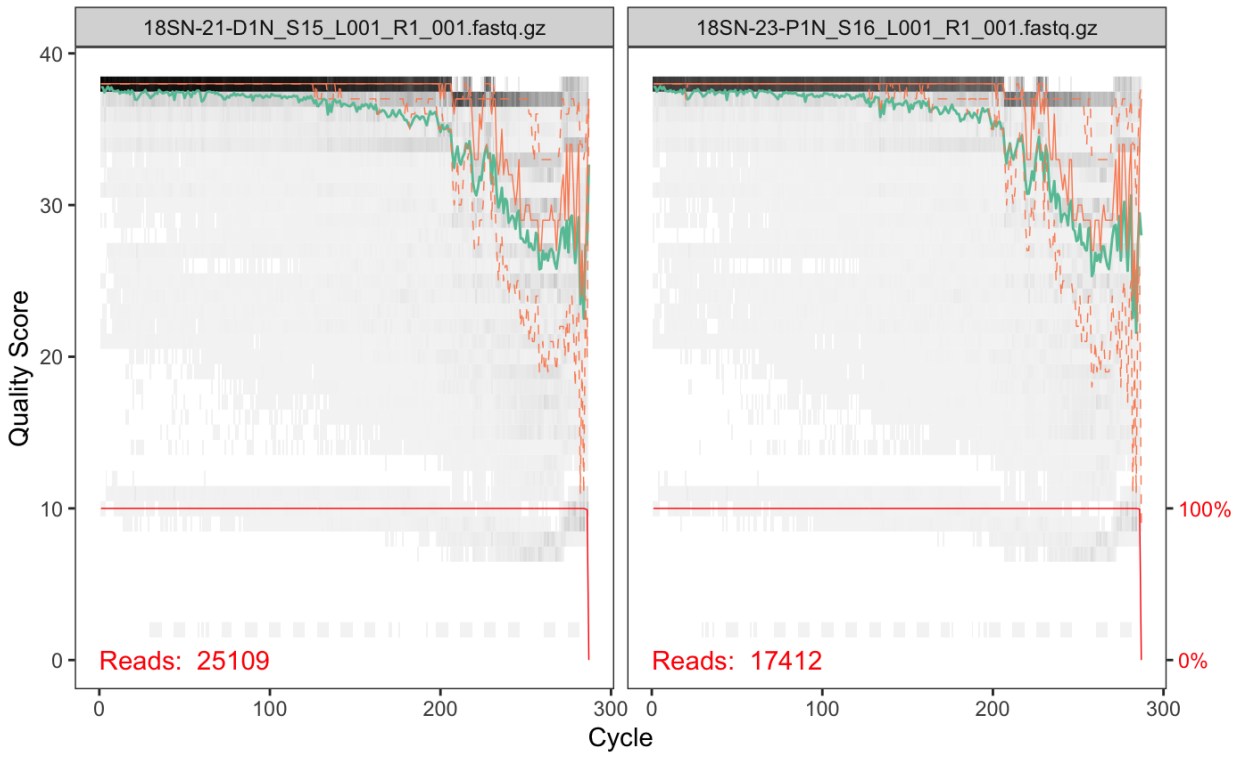
MgCl2 50 mM	1.25µL
10x buffer	2.5µL
Forward primer 10mM	1µL
Reverse primer 10mM	1µL
dNTPs 8mM	1.25µL
Apex Taq DNA polymerase	0.1875µL
Water	12.8125µL
DNA (sample)	2.5µL 20x diluted HW1 +2.5µL H2O, HW2, 2.5µL DW1, DW2 + 2.5 µL H2O; 5µL H1, H2, D1, D2, P1, P2

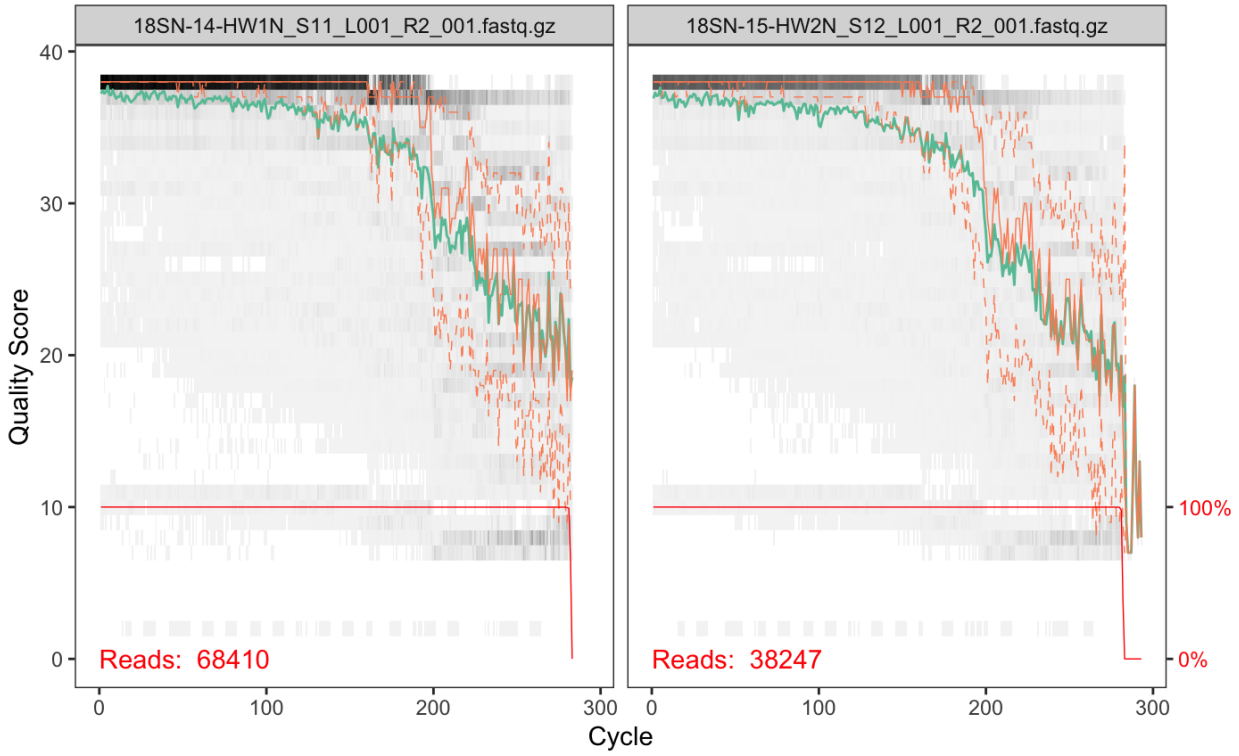
Thermocycler Protocol

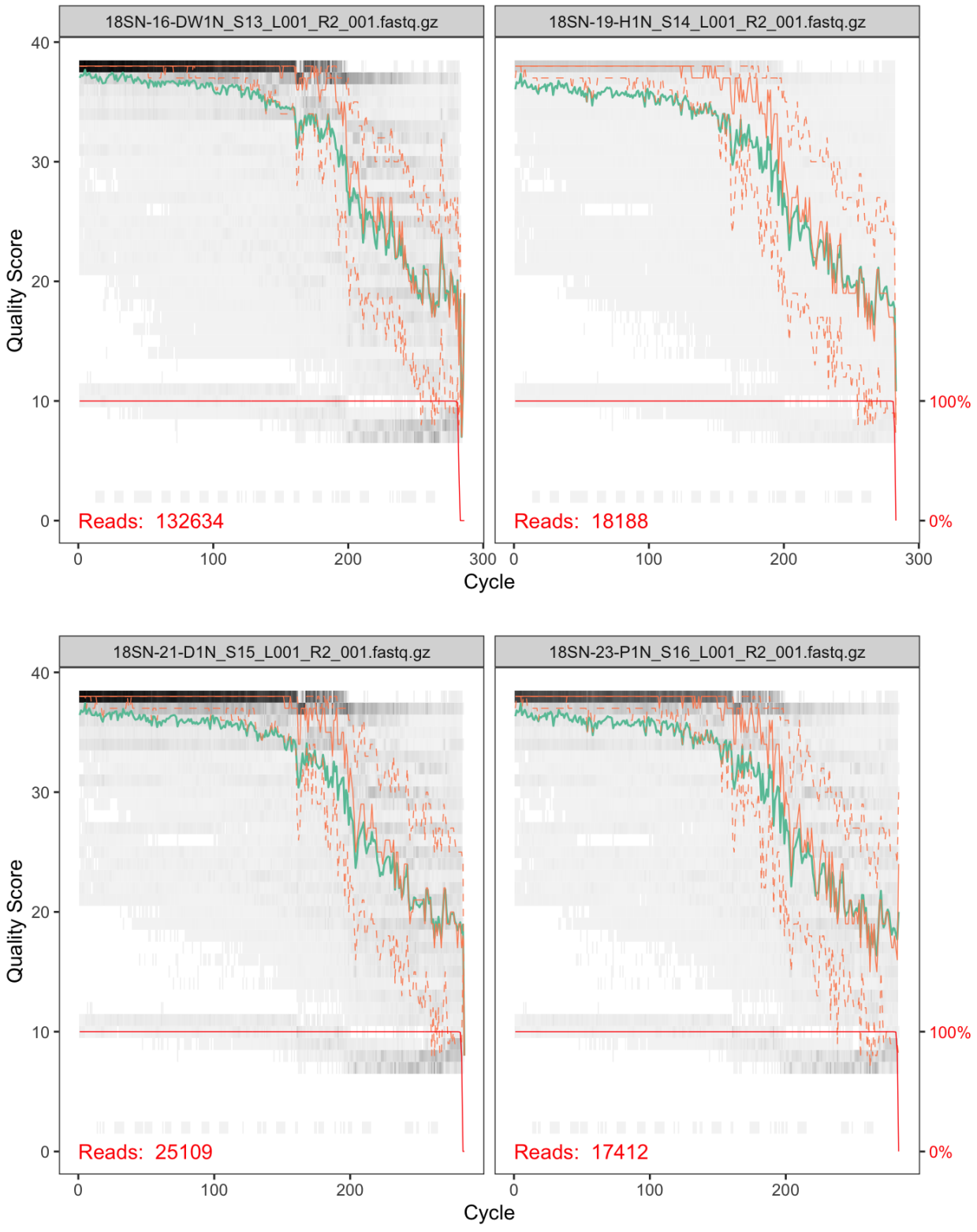
Temperature	**Time**	**Number of Cycles**
94°C	2 minutes	1
94°C	30 seconds	35
65°C	30 seconds	35
72°C	1 minute	35
72°C	10 minutes	1

Appendix III: Read Quality Profiles


18SN-14-HW1N_S11_L001_R1_001.fastq.gz
18SN-15-HW2N_S12_L001_R1_001.fastq.gz
18SN-16-DW1N_S13_L001_R1_001.fastq.gz
18SN-19-H1N_S14_L001_R1_001.fastq.gz
Quality Score
Cycle
Reads: 68410
Reads: 38247
Reads: 132634
Reads: 18188
100%
0%


18SN-21-D1N_S15_L001_R1_001.fastq.gz
18SN-23-P1N_S16_L001_R1_001.fastq.gz
Quality Score
Cycle
Reads: 25109
Reads: 17412
100%
0%

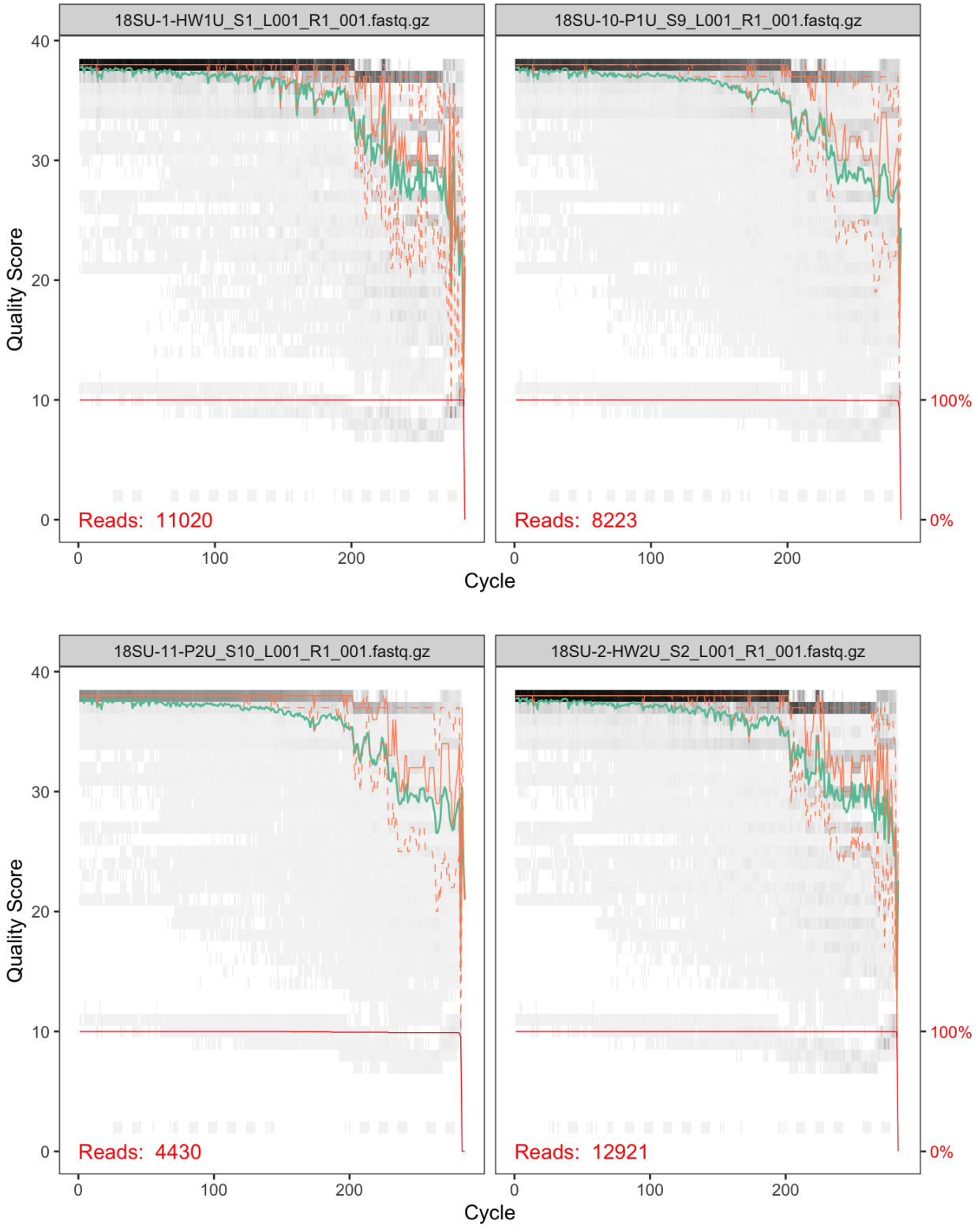

18SN-14-HW1N_S11_L001_R2_001.fastq.gz
18SN-15-HW2N_S12_L001_R2_001.fastq.gz
Quality Score
Cycle
Reads: 68410
Reads: 38247
100%
0%

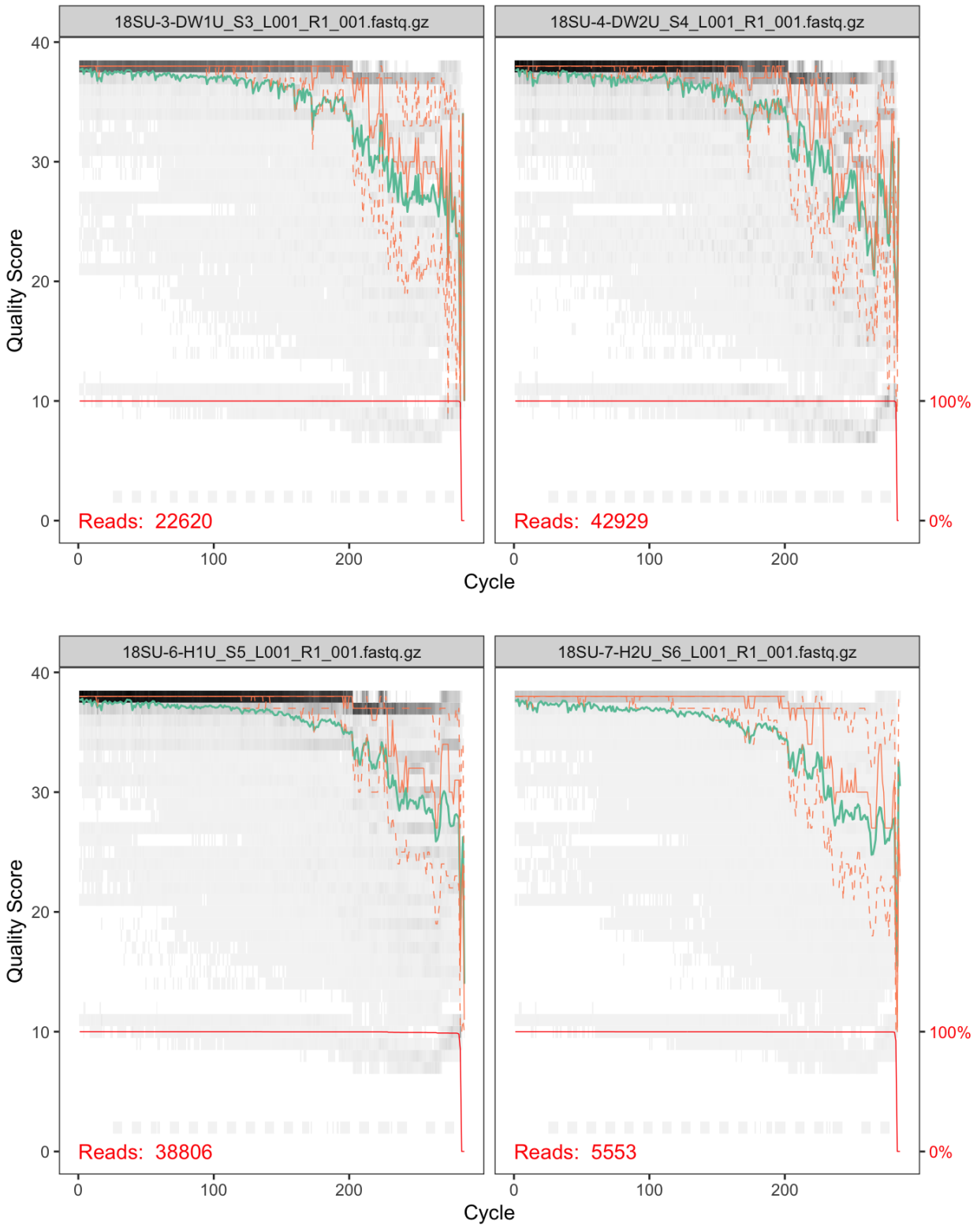


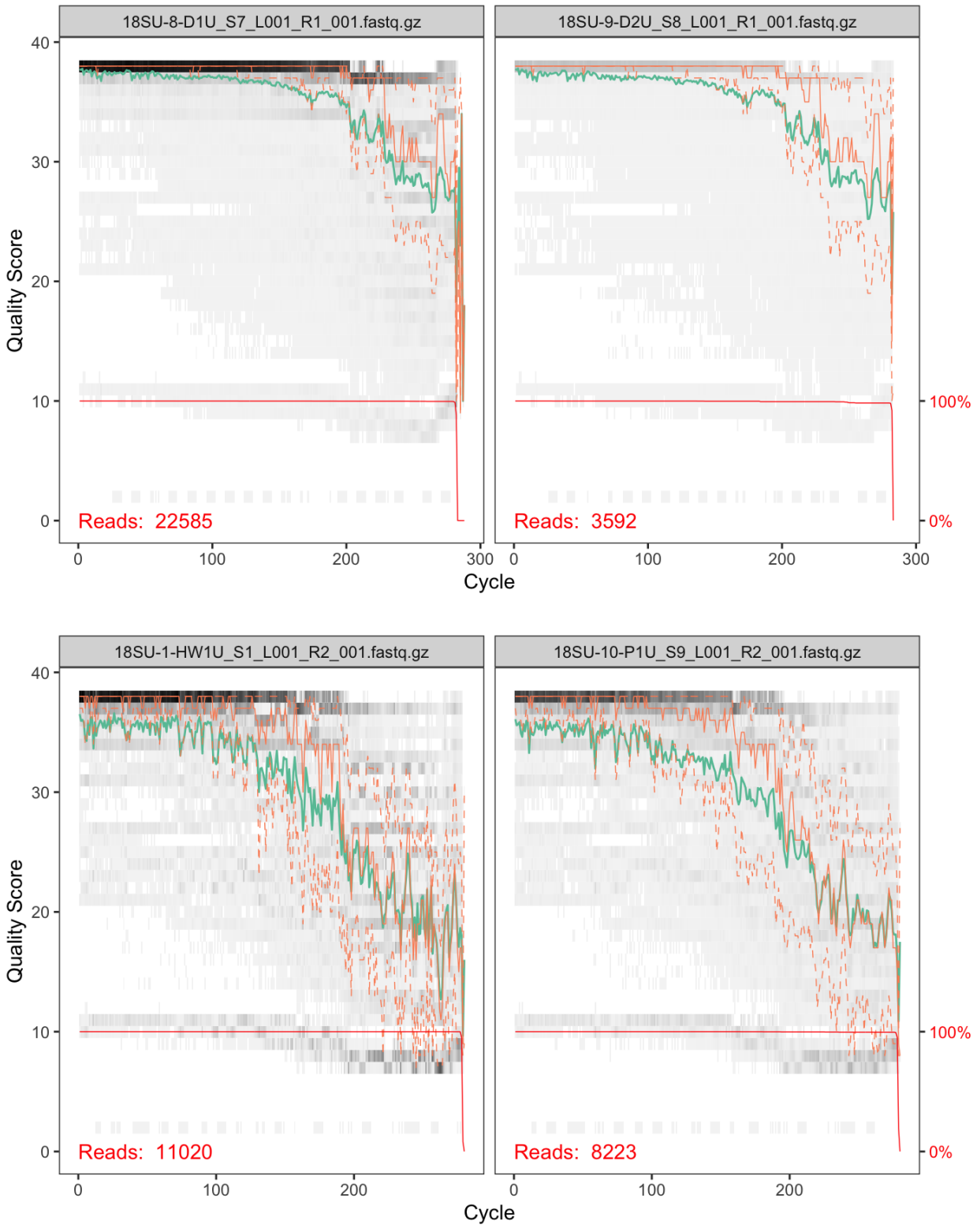
Appendix Figures 1-12. Read quality profiles of forward and reverse reads for non-metazoan primer samples. Quality score is assigned by the MiSeq and interpreted by dada2. X-axis

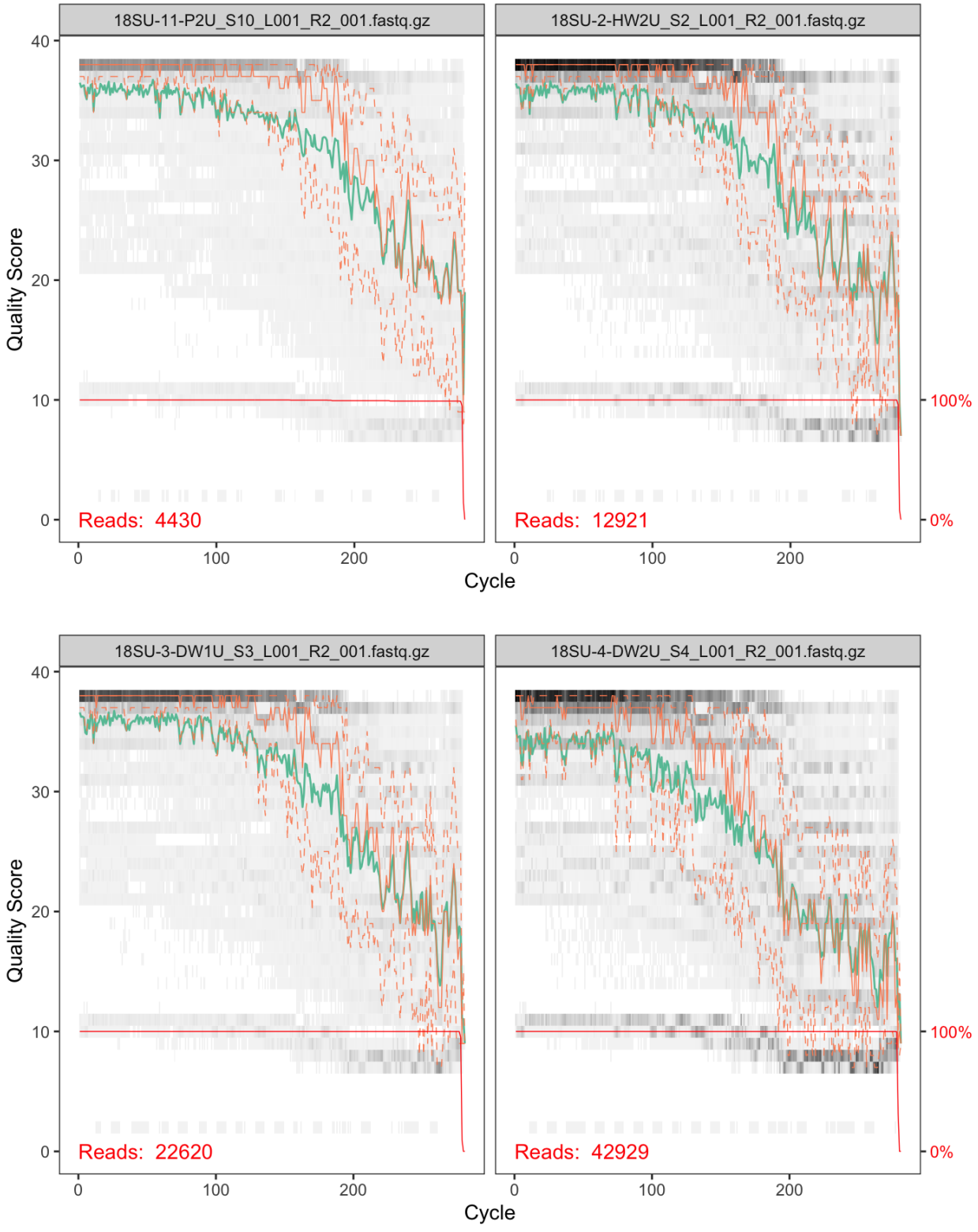
("Cycle") shows the length of the fragment in bp. Right y-axis shows the percentage of fragments reaching the indicated length. Number of reads is displayed in the lower left corner. Read quality score drop offs were used to determine where truncation would be carried out prior to merging paired sequences.

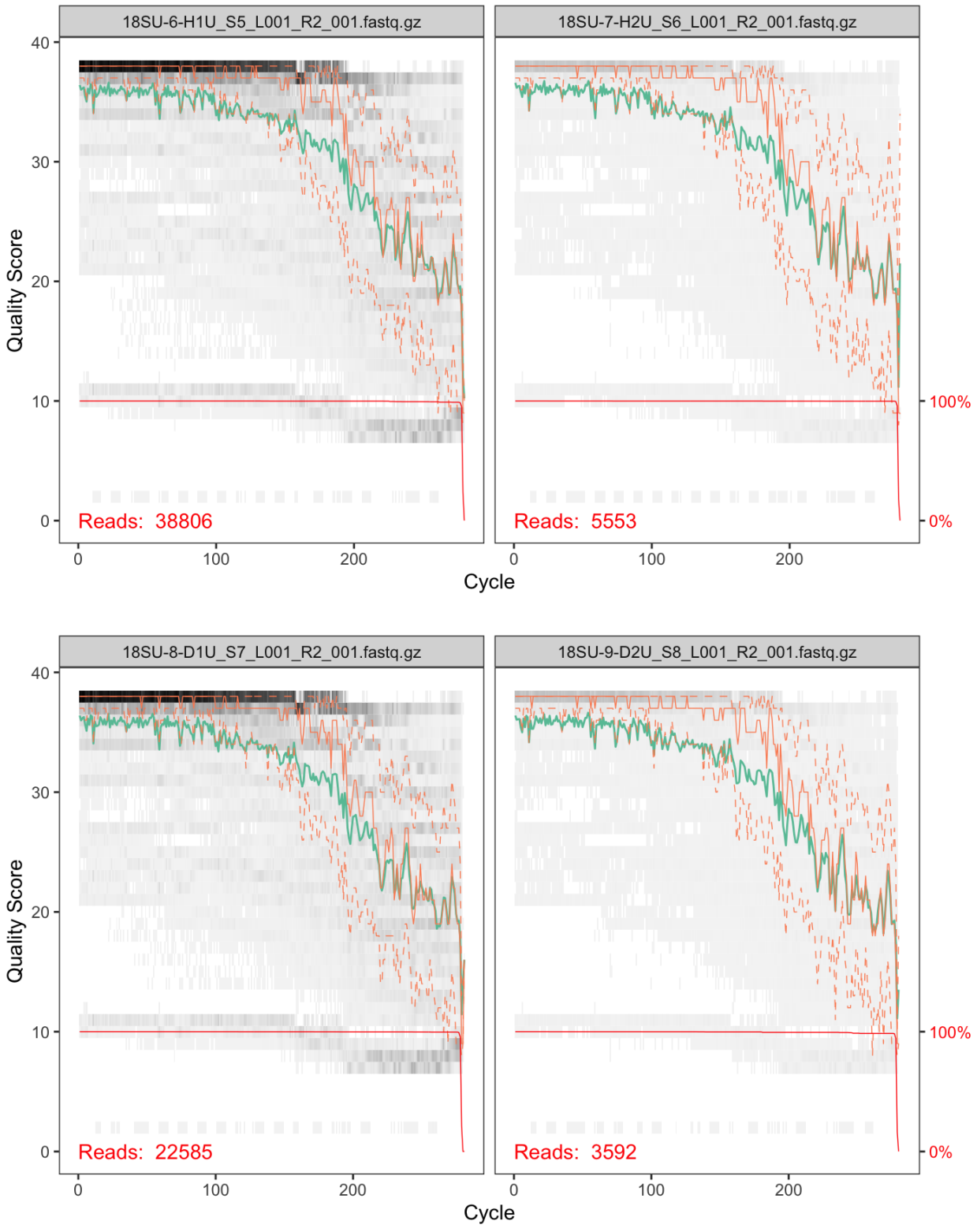
Universal:


18SU-1-HW1U_S1_L001_R1_001.fastq.gz
18SU-10-P1U_S9_L001_R1_001.fastq.gz
18SU-11-P2U_S10_L001_R1_001.fastq.gz
18SU-2-HW2U_S2_L001_R1_001.fastq.gz
Quality Score
Cycle
Reads: 11020
Reads: 8223
Reads: 4430
Reads: 12921
100%
0%
0
10
20
30
40
0
100
200


18SU-3-DW1U_S3_L001_R1_001.fastq.gz
18SU-4-DW2U_S4_L001_R1_001.fastq.gz
18SU-6-H1U_S5_L001_R1_001.fastq.gz
18SU-7-H2U_S6_L001_R1_001.fastq.gz
Quality Score
Cycle
Reads: 22620
Reads: 42929
Reads: 38806
Reads: 5553
100%
0%
0
10
20
30
40
100
200


18SU-8-D1U_S7_L001_R1_001.fastq.gz
18SU-9-D2U_S8_L001_R1_001.fastq.gz
18SU-1-HW1U_S1_L001_R2_001.fastq.gz
18SU-10-P1U_S9_L001_R2_001.fastq.gz
Quality Score
Cycle
Reads: 22585
Reads: 3592
Reads: 11020
Reads: 8223
100%
0%


18SU-11-P2U_S10_L001_R2_001.fastq.gz
18SU-2-HW2U_S2_L001_R2_001.fastq.gz
18SU-3-DW1U_S3_L001_R2_001.fastq.gz
18SU-4-DW2U_S4_L001_R2_001.fastq.gz
Quality Score
Cycle
Reads: 4430
Reads: 12921
Reads: 22620
Reads: 42929
100%
0%
0
10
20
30
40
0
100
200



Appendix Figures 13-23. Read quality profiles of forward and reverse reads for universal primer samples. Quality score is assigned by the MiSeq and interpreted by dada2. X-axis ("Cycle")

shows the length of the fragment in bp. Right y-axis shows the percentage of fragments reaching the indicated length. Number of reads is displayed in the lower left corner. Read quality score drop offs were used to determine where truncation would be carried out prior to merging paired sequences.