An enrichment approach for the recovery of viral and bacterial genomes from coral metagenomes

Bailey A. Wallace
University of Miami
Natascha S. Varona
University of Miami
Cynthia B. Silveira (cynthiasilveira@miami.edu)
University of Miami https://orcid.org/0000-0002-9533-5588

Method Article

Keywords: size fractionation, metagenome-assembled genomes, bacteriophage

Posted Date: January 17th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3867757/v1

Additional Declarations: The authors declare no competing interests.

Abstract

Difficulties in obtaining viral and bacterial genomes from corals have hindered the mechanistic understanding of these holobionts. Here, we introduce a size-fractionation approach to enrich viruses and bacteria from coral samples (tissues, mucus, and skeleton) for metagenome sequencing. Enriched metagenomes reduced host and Symbiodium DNA from 70 to 36% while increasing bacterial recovery by 9-fold. Remarkably, bacterial metagenome-assembled genomes were only recovered from the enriched metagenomes, and viral genome recovery increased by 3-fold, expanding the diversity of viruses detected. The high recovery of viral and bacterial genomes described here will facilitate the expansion of functional genomic studies in coral holobionts.

Background

Complex symbioses between the animal host, intracellular algae, protists, bacteria, archaea, and viruses are critical for the health and resilience of coral holobionts [1, 2]. While the mechanistic basis of coral-Symbiodinaceae relationships has been largely explored, the functions of other holobiont components remain elusive. For example, certain bacterial consortia were recently demonstrated to have a protective effect against bleaching at high temperatures [3], but how this protection is achieved is unknown. Among viruses, recent work showed that the induction of integrated prophages from bacterial genomes mediates competition within the holobiont, which, once disrupted, causes tissue loss [4]. One of the main limitations in pinpointing the functional roles of bacteria and viruses is the difficulty in obtaining bacterial and viral genomes from coral tissues. Host and Symbiodinaceae DNA overwhelmingly dominate shotgun metagenomes, preventing appropriate coverage of viral and bacterial genomes necessary for assembly. Most studies to date have relied on marker gene sequencing of bacteria (i.e., 16S rRNA) or methods for viral enrichment that incur severe compositional biases [5, 6, 7]. Here, we overcome this limitation by introducing a size fractionation method for enriching viral and bacterial genomes in shotgun metagenomic sequencing of coral samples.

Results and discussion

The Viral and Bacterial Enrichment (VBE; Fig. 1A) metagenomes had 36.39 ± 0.10% (mean, SE) coral or symbiont reads, compared to 69.83 ± 0.13% (mean, SE) in the controls, a 33.44% (1.92x) reduction in the VBE metagenomes (t-test, t(23) = -5.58, p = 1.124e-05) (Fig. 1B). Among quality-controlled reads, 2.72 ± 0.33% of VBE reads were bacterial, compared to only 0.30 ± 0.03% (mean, SE; here and hereafter) in controls, a 2.42% (9.01x) enrichment in VBE (t-test, t(23) = 4.12, p = 4.15e-04) (Fig. 1C). For coral and symbiont-filtered reads, VBE increased bacterial recovery by 5.02x (4.06 ± 0.09% and 0.81 ± 0.06% for VBE and control, respectively; t-test, t(23) = 4.29, p = 2.73e-04; Fig. 1D). Remarkably, high quality bacterial metagenome-assembled genomes (bMAGs) could only be generated from VBE metagenomes, with 13 total bMAGs with ≥ 50% completion and ≤ 10% contamination identified (Fig. 1E). VBE generated 543 putative viral contigs, compared to 76 from controls, which represented 0.74 ± 0.05% and 0.40 ± 0.05% of QC reads, respectively, a 1.87x increase for VBE (t-test, t(23) = 3.45, p = 2.17e-03) (Fig. 1F). Among coral

and symbiont-filtered reads, 0.46 ± 0.06% were viral in VBE and 0.16 ± 0.02% in controls, a 2.83x increase in viral recovery by VBE (t-test, t(23) = 2.6878, p = 1.31e-02; Fig. 1G). Bacterial and viral genome recovery through VBE was possible even at a relatively low sequencing coverage (37 million reads [5.5 Mbase] per sample on average). This recovery represents an improvement from previous studies that required coverages from 58 to 146 million reads [8, 9, 10, 11] to assemble bacterial MAGs and that were most successful with samples devoid of coral tissue, such as in studies focused on the skeleton [8].

Half of the viral genomes captured by VBE were classified as Duplodnaviria, as expected, given the DNA focus of the study, in addition to unclassified viral genomes (Fig. 2). Within Duplodnaviria, VBE recovered viral genomes belonging to nine families of bacteriophages. This contrasts with the control, where only unclassified genomes were recovered (Fig. 2). These results suggest that the VBE method incurred fewer compositional biases compared to previous methods [7, 12, 13, 14, 15]. VBE yielded enough coverage for viral genome assembly, unlike previous unbiased sequencing approaches that had to limit analyses to individual reads due to the difficulty of assembling, compromising functional interpretations [16, 17].

VBE uses a coral fragment of $1cm^2$, not much larger than a parrotfish bite, and is minimally invasive to corals. Therefore, this method could be easily adapted for the recovery of bacterial and viral RNA, which would presumably require larger amounts of input material [18]. This method pooled all holobiont compartments (skeleton, tissue, and mucus), and adaptations are required for application to specific compartments [9, 19]. Increasing the size of the coral input sample and sequencing depth and incorporating long-read data may lead to the resolution of more complete genomes covering a larger biological diversity.

Conclusion

The VBE method introduced here significantly improved bacterial and viral genome assembly from coral shotgun metagenomes. VBE avoided using chloroform, amplification, or density gradients, capturing a larger diversity of viruses with low compositional bias. VBE will facilitate the expansion of functional genomic studies in corals and can be readily adapted to other holobionts.

Methods

A step-by-step description of the VBE protocol is publicly available on protocols.io [20]. Fragments of the coral *Orbicella faveolata* (N = 28) were collected via SCUBA in July 2021 along the southwestern coast of Curaçao in the Caribbean (Table S1). Specimens of approximately $1cm^2$ (containing mucus, tissue, and skeleton) were collected using a chisel and hammer and placed in Ziplock polyethylene bags. The specimens were briefly placed on ice during transfer to the laboratory at the CARMABI research station, where they were flash-frozen and stored at − 80°C until later processing. The samples were subjected to two DNA extraction protocols for comparison: 19 samples were processed using the Viral and Bacterial Enrichment (VBE) method developed here, and 6 underwent bulk extraction and sequencing and were treated as controls. For both extraction methods, coral fragments were thawed, crushed to a fine gravel

texture using a sterile mortar and pestle, and suspended in 150 µL of sterile artificial seawater. Control coral homogenates were extracted using a DNeasy PowerSoil Kit (QIAGEN, Germantown, MD, United States), modified with the addition of 20 µL of Proteinase K incubated at 56°C for 10 minutes prior to the kit's lysis steps. For VBE, the coral homogenate was placed into a tube containing ~ 0.2 g of 425–600 µm sterile glass beads (Sigma-Aldrich, St. Louis, MO, United States) and vortexed at speed 3 for 5 minutes (VWR Analog Vortex Mixer; VWR, Radnor, PA, United States). The supernatant was transferred to a clean tube, brought up to 1 mL with sterile ASW, and treated with DNAse I (20 U/mL final concentration) in DNAse I buffer (Invitrogen, Waltham, MA, United States) at 25°C for 2 hours. DNAse activity was stopped with EDTA (0.005 M final concentration; Genesee Scientific, San Diego, CA, United States), and the sample was filtered through an 8.0 µm membrane (Cytiva, Marlborough, MA, United States). The flowthrough was collected, transferred to a 100 KDa Amicon Centrifugal Unit (Sigma-Aldrich), and centrifuged at 3,200 g for 30 minutes. After concentration, each side of the Amicon filter was rinsed and incubated at 56°C for 1 hour (per side) in 200 µL Buffer T1 and 20 µL Proteinase K from the NucleoSpin® Tissue Kit (MACHEREY-NAGEL Inc., Allentown, PA, United States) and DNA extraction proceeded from step 3 of the kit. DNA was eluted in 100 µL of PCR-grade water and quantified with a Qubit 2.0 Fluorometer (Invitrogen). Library preparation and whole genome sequencing were conducted by Azenta Life Sciences (South Plainfield, NJ, United States). Metagenomic libraries were generated with the NEBNext Ultra DNA Library Preparation kit following the manufacturer's instructions (New England Biolabs, Ipswich, MA, United States) and paired-end sequenced (2 x 150 bp) on an Illumina HiSeq 4000 platform (Illumina, San Diego, CA, United States).

Raw metagenomic reads were adapter-trimmed, quality-filtered (trimq = 30, maq = 30), and entropy-filtered (entropy = 0.90) using BBDuk [21] generating 594,487,714 quality-controlled reads. To remove coral host and Symbiodinaceae reads, the quality-controlled (QC) reads were mapped to the genome of the coral host, *O. faveolata* (GCA_001896105.1), and Symbiodinaceae (Clade A) (GCA_003297005.1) using Bowtie2 (–mp 4 and -X 1000) [22]. Mapped reads were removed using SAMtools v1.18 [23] and quantified with SeqKit [24], yielding 336,288,370 coral and Symbiodinaceae filtered reads (Table S1). Quality-controlled and filtered reads from each sample were assembled separately using metaSPAdes v3.15.5 using base parameters [25].

Bacterial reads were identified with Kaiju v1.9.0 using the Progenomes database [26] and normalized by the number of quality or filtered reads in each sample to obtain the percentage of bacterial reads in each metagenome. Bacterial and archaeal MAGs were generated by integrating the single-sample coverage binning outputs from MaxBin2 v2.2.7 [27], MetaBat2 v2.15 [28], and CONCOCT v1.1.0 [29]. The resulting bins were consolidated and improved with the metaWRAP v1.2.1 [30] bin refinement module. The quality of refined bins was assessed with CheckM2 v1.0.2 [31] to select bins with ≥ 50% completion and ≤ 10% contamination. Bins with low-confidence predictions were removed. VIBRANT v1.2.1 identified putative viral genome fragments among contigs [32]. Viral contigs were pooled by DNA extraction method to create two databases, "VBE" and "Control". The number of reads mapped to their respective databases at 80% identity with SMALT v0.7.6 [33] was normalized by the total number of quality and filtered reads and by the length of each contig [34] to obtain the viral fractional abundance. Comparisons between VBE and

controls were based on the Student's t-test between the two groups, apart from Fig. 1E, which displays a simple count of bMAGs generated from each group. Taxonomic assignment of phages was performed with the Phage Taxonomy Tool (PTT) [35], with the reference taxonomy edited to reflect updated ICTV classifications [36].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Metagenomes generated here are publicly available through the NCBI SRA BioProject: PRJNA1061506. Codes utilized for the analyses are available through Github (https://github.com/Silveira-Lab/Wallace_Coral_Holobiont_Viruses/blob/main/Coral_Metagenomes.md) A step-by-step description of the protocol is available on Protocols.io [20] (https://www.protocols.io/view/virus-and-prokaryote-enrichment-in-coral-dna-metag-q26g7p1y3gwz/v1).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the University of Miami Provost Research Award (UM PRA 2022-2547 to CS). BW was funded by the NSF GRFP (Fellow ID: 2023353157), NV was funded by the Maytag Fellowship (Grant ID: PG015171), and all three authors were partially funded by the NASA Exobiology Program (80NSSC23K0676 to CS).

Authors' contributions

CS conceptualized and led the development of the method. BW and NV conducted sampling, method development, sample processing, and analyses. BW wrote the first draft, and NV and CS edited the manuscript.

Acknowledgments

We thank Mark Vermeij and the CARMABI research station in Curaçao for field support. We also thank Andrew Baker and Kathleen Sealey for access to coral specimens for the first trials in method development.

References

1. Knowlton N, Rohwer F. Multispecies microbial mutualisms on coral reefs: the host as a habitat. Am Nat. 2003 Oct;162(S4):S51-62.
2. Voolstra CR, Suggett DJ, Peixoto RS, Parkinson JE, Quigley KM, Silveira CB, et al. Extending the natural adaptive capacity of coral holobionts. Nat Rev Earth Environ. 2021 Oct 12;2(11):747–62.
3. Santoro EP, Borges RM, Espinoza JL, Freire M, Messias CSMA, Villela HDM, et al. Coral microbiome manipulation elicits metabolic and genetic restructuring to mitigate heat stress and evade mortality. Sci Adv. 2021 Aug;7(33):eabg3088.
4. Wang W, Tang K, Wang P, Zeng Z, Xu T, Zhan W, et al. The coral pathogen *Vibrio coralliilyticus* kills non-pathogenic holobiont competitors by triggering prophage induction. Nat Ecol Evol. 2022 Jun 30;6(8):1132–44.
5. Gignoux-Wolfsohn SA, Precht WF, Peters EC, Gintert BE, Kaufman LS. Ecology, histopathology, and microbial ecology of a white-band disease outbreak in the threatened staghorn coral *Acropora cervicornis*. Diseases of Aquatic Organisms. 2020 Jan 30;137(3):217-37
6. Rosales SM, Huebner LK, Evans JS, Apprill A, Baker AC, Becker CC, Bellantuono AJ, Brandt ME, Clark AS, Del Campo J, Dennison CE. A meta-analysis of the stony coral tissue loss disease microbiome finds key bacteria in unaffected and lesion tissue in diseased colonies. ISME communications. 2023 Mar 9;3(1):19.
7. Weynberg KD, Wood-Charlson EM, Suttle CA, van Oppen MJ. Generating viral metagenomes from the coral holobiont. Frontiers in Microbiology. 2014 May 7;5:206.
8. Cárdenas A, Raina JB, Pogoreutz C, Rädecker N, Bougoure J, Guagliardo P, Pernice M, Voolstra CR. Greater functional diversity and redundancy of coral endolithic microbiomes align with lower coral bleaching susceptibility. The ISME journal. 2022 Oct;16(10):2406-20.
9. Rosales SM, Huebner LK, Clark AS, McMinds R, Ruzicka RR, Muller EM. Bacterial metabolic potential and micro-eukaryotes enriched in stony coral tissue loss disease lesions. Frontiers in Marine Science. 2022 Jan 6;8:776859.
10. Robbins SJ, Singleton CM, Chan CX, Messer LF, Geers AU, Ying H, Baker A, Bell SC, Morrow KM, Ragan MA, Miller DJ. A genomic view of the reef-building coral Porites lutea and its microbial symbionts. Nature Microbiology. 2019 Dec;4(12):2090-100.
11. Sun F, Yang H, Wang G, Shi Q. Combination analysis of metatranscriptome and metagenome reveal the composition and functional response of coral symbionts to bleaching during an El Niño Event. Frontiers in Microbiology. 2020 Mar 20;11:448.
12. Ackermann, HW. The Bacteriophages. 2nd ed. Oxford University Press; 2006. Chapter 2, Classification of bacteriophages; p.8-16.
13. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Sci Rep. 2015 Nov 12;5:16532.

14. Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC genomics. 2015 Dec;16:1-5.
15. Karlsson OE, Belák S, Granberg F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. Biosecurity and bioterrorism: biodefense strategy, practice, and science. 2013 Sep 1;11(S1):S227-34.
16. Cárdenas A, Ye J, Ziegler M, Payet JP, McMinds R, Vega Thurber R, et al. Coral-associated viral assemblages from the Central Red Sea align with host species and contribute to holobiont genetic diversity. Frontiers in Microbiology. 2020 Sep 30;11:572534.
17. Messyasz A, Rosales SM, Mueller RS, Sawyer T, Correa A, Thurber AR, Vega Thurber R. Coral bleaching phenotypes associated with differential abundances of nucleocytoplasmic large DNA viruses. Frontiers in Marine Science. 2020:789.
18. Levin RA, Voolstra CR, Weynberg KD, Van Oppen MJ. Evidence for a role of viruses in the thermal sensitivity of coral photosymbionts. The ISME Journal. 2017 Mar;11(3):808-12.
19. Lima LF, Weissman M, Reed M, Papudeshi B, Alker AT, Morris MM, et al. Modeling of the coral microbiome: the influence of temperature and microbial network. MBio. 2020 Apr 28;11(2):10-128.
20. Varona N, Wallace B, Silveira S. Virus and prokaryote enrichment in coral DNA metagenomes. 2023 Sep 21; protocols.io. Available from: https://dx.doi.org/10.17504/protocols.io.q26g7p1y3gwz/v1
21. Bushnell, Brian. BBMap short read aligner, and other bioinformatic tools. Available from: https://sourceforge.net/projects/bbmap/Accessed: [November 21, 2022]
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Apr;9(4):357–9.
23. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.
24. Shen W, Le S, Li Y, Hu F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLOS ONE. 2016 Oct 5;11(10):e0163962.
25. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017 May;27(5):824–34.
26. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun. 2016 Apr 13;7(1):11257.
27. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics. 2016 Feb 15;32(4):605–7.
28. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019 Jul 26;7:e7359.
29. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. Nature methods. 2014 Nov;11(11):1144-6.

30. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018 Sep 15;6(1):158.
31. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods. 2023 Aug;20(8):1203–12.
32. Kieft K, Zhou Z, Anantharaman K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome. 2020 Jun 10;8(1):90.
33. Wellcome Sanger Institute. SMALT, A mapper for DNA sequencing reads. Available from: https://sourceforge.net/projects/smalt/. Accessed: [March 15, 2023]
34. Cobián Güemes AG, Youle M, Cantú VA, Felts B, Nulton J, Rohwer F. Viruses as Winners in the Game of Life. Annual Review of Virology. 2016;3(1):197–214.
35. Kieft K, Zhou Z, Anderson RE, Buchan A, Campbell BJ, Hallam SJ, Hess M, Sullivan MB, Walsh DA, Roux S, Anantharaman K. Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. Nature communications. 2021 Jun 9;12(1):3503.
36. Turner D, Shkoporov AN, Lood C, Millard AD, Dutilh BE, Alfenas-Zerbini P, van Zyl LJ, Aziz RK, Oksanen HM, Poranen MM, Kropinski AM. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. Archives of Virology. 2023 Feb;168(2):74.

Supplementary Table

Table S1 is not available with this version

Figures

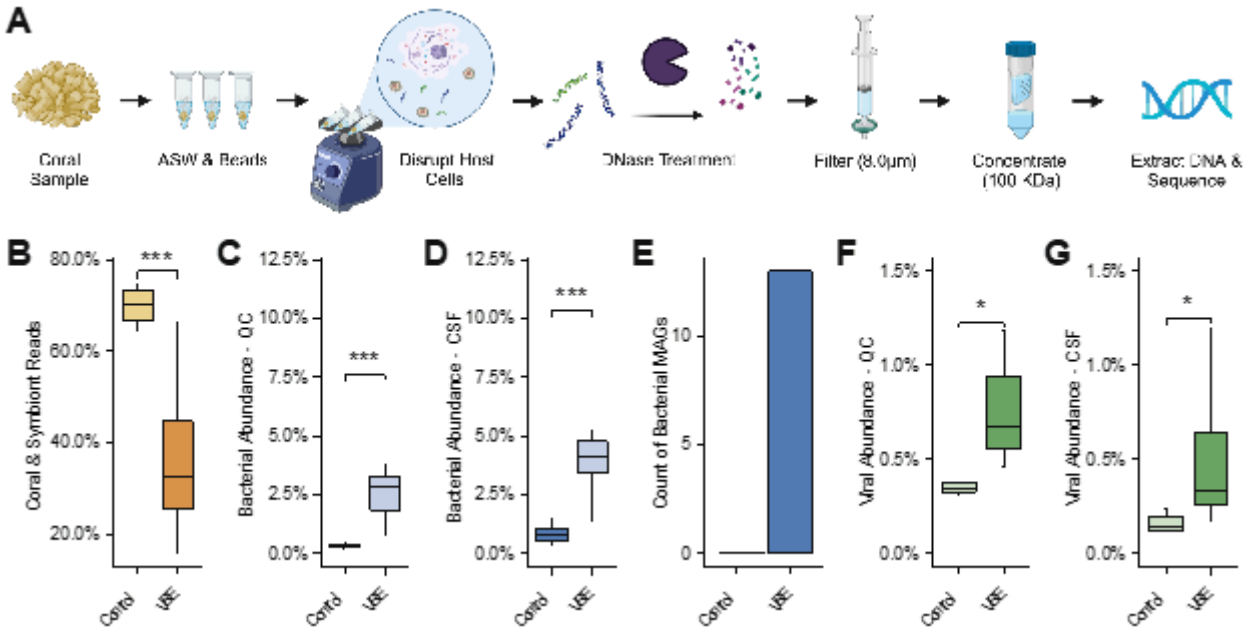

Figure 1

Enrichment of viruses and bacteria from coral samples. (A) Simplified workflow depicting the VBE method for enrichment of viruses and bacteria in coral metagenomes. **(B-G)** Recovery of coral and symbiont (orange), bacteria (blue), and viruses (green) in VBE and control metagenomes. **(B)** Percent of coral and symbiont reads, **(C)** bacterial abundance within quality-controlled (QC) reads, **(D)** Bacterial abundance within coral and symbiont-filtered (CSF) reads, **(E)** Number of bacterial MAGs (the bar plot indicates the counts, not relative abundances), **(F)** Viral fractional abundance in QC reads and **(G)** in CSF reads. Stars represent p-values from Student's t-test in panels B, C, D, F, and G (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Error bars represent the standard error.

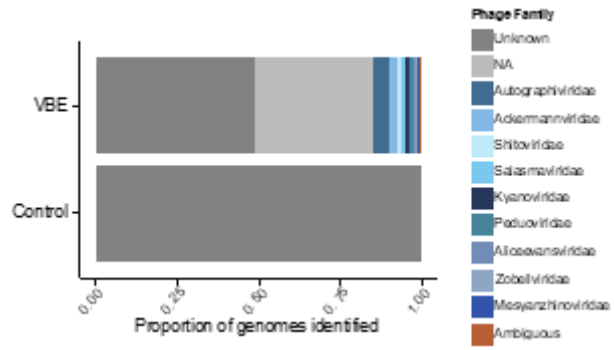

Figure 2

Taxonomy of viral genomes captured by the Viral and Bacterial Enrichment (VBE) method. Taxonomic assignments of bacteriophages in the VBE and Control groups. Dark gray indicates genomes without any taxonomic classification (unknown), while light gray (NA) indicates genomes classified as Duplodnaviria but without family-level classification.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx