



Comparative Omics and Trait Analyses of Marine *Pseudoalteromonas* Phages Advance the Phage OTU Concept

Melissa B. Duhaime^{1*}, Natalie Solonenko², Simon Roux², Nathan C. Verberkmoes³, Antje Wichels⁴ and Matthew B. Sullivan^{2,5*}

¹ Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, United States, ² Department of Microbiology, The Ohio State University, Columbus, OH, United States, ³ Department of Biological Sciences, Border Biomedical Research Center, University of Texas at El Paso, El Paso, TX, United States, ⁴ Biologische Anstalt Helgoland, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Helgoland, Germany, ⁵ Department of Civil, Environmental, and Geodetic Engineering, The Ohio State University, Columbus, OH, United States

OPEN ACCESS

Edited by:

William Michael McShan,
University of Oklahoma Health
Sciences Center, United States

Reviewed by:

Claire Bertelli,
Simon Fraser University, Canada
Alison Buchan,
University of Tennessee, Knoxville,
United States

*Correspondence:

Melissa B. Duhaime
duhaimem@umich.edu
Matthew B. Sullivan
mbsulli@gmail.com

Specialty section:

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

Received: 24 February 2017

Accepted: 20 June 2017

Published: 06 July 2017

Citation:

Duhaime MB, Solonenko N, Roux S, Verberkmoes NC, Wichels A and Sullivan MB (2017) Comparative Omics and Trait Analyses of Marine *Pseudoalteromonas* Phages Advance the Phage OTU Concept. *Front. Microbiol.* 8:1241. doi: 10.3389/fmicb.2017.01241

Viruses influence the ecology and evolutionary trajectory of microbial communities. Yet our understanding of their roles in ecosystems is limited by the paucity of model systems available for hypothesis generation and testing. Further, virology is limited by the lack of a broadly accepted conceptual framework to classify viral diversity into evolutionary and ecologically cohesive units. Here, we introduce genomes, structural proteomes, and quantitative host range data for eight *Pseudoalteromonas* phages isolated from Helgoland (North Sea, Germany) and use these data to advance a genome-based viral operational taxonomic unit (OTU) definition. These viruses represent five new genera and inform 498 unaffiliated or unannotated protein clusters (PCs) from global virus metagenomes. In a comparison of previously sequenced *Pseudoalteromonas* phage isolates ($n = 7$) and predicted prophages ($n = 31$), the eight phages are unique. They share a genus with only one other isolate, *Pseudoalteromonas* podophage RIO-1 (East Sea, South Korea) and two *Pseudoalteromonas* prophages. Mass-spectrometry of purified viral particles identified 12–20 structural proteins per phage. When combined with 3-D structural predictions, these data led to the functional characterization of five previously unidentified major capsid proteins. Protein functional predictions revealed mechanisms for hijacking host metabolism and resources. Further, they uncovered a hybrid siphon-myovirus that encodes genes for Mu-like infection rarely described in ocean systems. Finally, we used these data to evaluate a recently introduced definition for virus populations that requires members of the same population to have >95% average nucleotide identity across at least 80% of their genes. Using physiological traits and genomics, we proposed a conceptual model for a viral OTU definition that captures evolutionarily cohesive and ecologically distinct units. In this trait-based framework, sensitive hosts are considered viral niches, while host ranges and infection efficiencies are tracked as viral traits. Quantitative host range assays revealed conserved traits within virus OTUs that break down between OTUs, suggesting the defined units capture

niche and fitness differentiation. Together these analyses provide a foundation for model system-based hypothesis testing that will improve our understanding of marine copiotrophs, as well as phage–host interactions on the ocean particles and aggregates where *Pseudoalteromonas* thrive.

Keywords: phage, marine microbiology, particle-associated, comparative genomics, *Pseudoalteromonas*

INTRODUCTION

Microbes and their metabolic outputs impact diverse ecosystem functions (Falkowski et al., 2008) and viruses tune these microbial metabolisms through mortality, horizontal gene transfer, and host metabolic reprogramming (Fuhrman, 1999; Wommack and Colwell, 2000; Suttle, 2007; Breitbart, 2012; Brum and Sullivan, 2015). In the oceans, micron- to millimeter-sized particles (e.g., marine snow, fecal pellets, gelatinous exudates, and carcasses of zooplankton) are hotspots of this microbial metabolic activity (Azam, 1998), in particular heterotrophic activity (Ploug et al., 1999). The microbial processes that take place on and in ocean particles modulate the flux of organic matter, thereby impacting the efficiency of the biological carbon pump and ultimately the planetary climate system (Shanks and Trent, 1980; Ducklow et al., 2001). Viruses are present and active on these sinking particles (Bratbak et al., 1990; Proctor and Fuhrman, 1991; Weinbauer et al., 2009). Recently, depth-stratified viral community sequencing studies (viromes) have provided “genetic tracer” evidence that surface water viruses sink to the deep sea, which has been hypothesized to occur in association with particles (Brum and Sullivan, 2015; Brum et al., 2015; Hurwitz et al., 2015). Following on this, a recent study that examined the coupling between omics-based plankton community structure and surface ocean carbon export found that viruses are the best predictors of ocean carbon flux (Guidi et al., 2016). These recent reports highlight the need to revisit decades-old hypotheses about particle-adsorbed viruses (Proctor and Fuhrman, 1991). Further, they invite questions, such as what is the impact of sinking particle microcosms on virus and host biogeography, phage–host co-evolution, and predator–prey infection dynamics? And in turn, what are the impacts of these viral processes on the sinking rates of ocean particles and thereby the efficiency of the global biological carbon pump?

Pseudoalteromonas spp. (Gammaproteobacteria; Alteromonadales) are ideal hosts for developing a model particle-associated phage and host system. On the global scale, *Pseudoalteromonas* spp. are ubiquitous marine heterotrophs adapted to life on particles (Thomas et al., 2008), where they are highly represented and commonly constitute up to 20% of particle-associated (Fontanez et al., 2015) and particle-enriched (Smriga et al., 2016) microbial communities. Further, *Pseudoalteromonas* is the heterotrophic genus most strongly correlated with carbon export in the world’s oceans (Guidi et al., 2016). On the microscale, members of the genus have shown a strong and rapid chemotactic response toward dissolved organic matter plumes (Stocker et al., 2008), implicating them as model gradient-responding copiotrophs for the study of ocean particle ecology. Relationships with viruses are likely

to impact the role of *Pseudoalteromonas* in the ocean particle habitat. In addition to the direct impact of host cell lysis on carbon flux, *Pseudoalteromonas* phages may impact the behavioral ecology of their microbial hosts in a manner that could modulate the magnitude of this effect: a filamentous *Pseudoalteromonas* phage has been shown to enhance the motility and chemotaxis of its infected host in culture (Yu et al., 2015).

Currently, phages with >95% average nucleotide identity (ANI) across at least 80% of their genes are assigned to a single phage population (Brum et al., 2015; Gregory et al., 2016), or as termed here, “phage OTU”—consistent with the 95% ANI cut-off proposed for microbial species (Konstantinidis and Tiedje, 2005). For phages, this concept has been in development for some time and has been supported by the stable spatial and temporal distribution of T4-like cyanomyophage isolates based on gene marker (Marston and Amrich, 2009) and full-genome analyses (Marston and Martiny, 2016), as well as genome-wide analyses of viral-tagged metagenomic contigs from wild virus populations (Deng et al., 2014). Most recently, a phylogenomic analysis of 142 marine T4-like cyanophages isolated on a single host (*Synechococcus* WH7803) observed that when this >95% threshold was applied, (i) recombination rates were greater within genotypic phage populations than between them, indicative of intra-population barriers to gene flow and (ii) different genes were under selection in the different populations, while the selection profile was conserved within populations (Gregory et al., 2016). While trends have emerged from genomic data to support a phage operational taxonomic unit (OTU) definition, these studies have not yet included physiological tests that would enable evaluation of fitness differences between genotypic populations.

To inform our understanding of the role of *Pseudoalteromonas* (PSA) phages in particle and ocean ecology, we sought to characterize the life history traits (e.g., host range, burst size, latent period), structural proteomes, and genomes of eight newly introduced PSA phages isolated offshore of the island Helgoland in the North Sea. These combined trait-based and comparative genomics analyses allowed us to test the validity of proposed sequence-based virus population delineations described above (Brum et al., 2015; Gregory et al., 2016). To do so, we considered *Pseudoalteromonas* hosts as viral niches and measured infection traits indicative of viral fitness in each niche. We hypothesized that if the currently proposed population (or phage OTU) definition were meaningful (i.e., capture evolutionarily and ecologically cohesive populations; Polz et al., 2006), trait-based differentiation would arise consistent with these genome-based boundaries. These analyses provide a baseline for understanding the ecological and evolutionary impact of viruses

infecting *Pseudoalteromonas*, a model ocean particle-associated copiotrophic microbe.

MATERIALS AND METHODS

Phage Harvesting and DNA Extraction

Phage and hosts were isolated from unfiltered, whole seawater in 1990 from the 'Kabeltonne' station near Helgoland, Germany (54.18 N 7.9 E), 55 km north of the German coast in the North Sea (Moebus, 1992; Wichels et al., 1998). In 2010, phages were recovered from liquid lysates stored at room temperature and hosts from glycerol stocks stored in liquid nitrogen. Host growth and infections were performed as described (Duhaime et al., 2011), with the exception of host growth media, which was altered to a "*Pseudoalteromonas* Zobell Media (PZM)." PZM was comprised of 50% of all components of the previously described Zobell Media, except sea salts remain at 25 g L⁻¹. Host cultures were grown at 21°C and shaken at 150 rpm. Phage DNA was extracted using the Wizard PCR Prep DNA Purification Resin and Mini-columns (Promega, San Luis Obispo, CA, United States) per manufacturer provided protocol. Phage genomes, annotations, and associated metadata are publically available in the Joint Genome Institute's Integrated Microbial Genomes (JGI-IMG) database with the following IMG Taxon IDs: 2582581227, 2582581228, 2582581229, 2582581230, 2582581231, 2582581232, 2582581235, 2622736497.

Transmission Electron Microscopy

Concentrated viral lysates (>10⁸ viruses ml⁻¹) were CsCl-purified (Duhaime et al., 2011) and 5 µl deposited on formvar coated 200 mesh copper grids (Electron Microscopy Sciences, Hatfield, PA, United States) that had been glow discharged for 3 min with a sputter coater (Hummer 6.2, Anatech, Union City, CA, United States). Grids were then stained with three drops of 0.02 µm-filtered 2% (w/v) uranyl acetate and for 30 s followed by three 10-s washes in ultra-pure water. All liquid was wicked away with filter paper to achieve negatively stained viral specimen. Grids were left to dry overnight in a desiccator at ambient temperature. Dry grids were visualized with a transmission electron microscope (Philips CM12, FEI, Hillsboro, OR, United States) at 80 kV accelerating voltage and 65,000–100,000 magnification. Micrographic images were collected using a Macrofire Monochrome CCD camera (Optronics, Goleta, CA, United States).

Phage Infection Properties: One-Steps, Burst Sizes, Latent Periods

Before the one-step experiment was performed, the relationship between culture optical density (OD) and colony-forming units was established through a host growth curve. One-step experiments were performed in triplicate at targeted initial multiplicities of infection (MOI) of 0.1. Viruses were added to host cultures in mid-exponential (log linear) growth phase, whereby the host cell concentration was determined according to the OD-CFU correlation and the viruses were added accordingly

to achieve the desired MOI (typically 10⁷ viruses added to 10⁸ host cells). No-phage controls were monitored in parallel. At $t = 0$, 15 min post-inoculation with viruses, cultures were diluted 1:100 in 50 ml PZM in a 250 ml glass Erlenmeyer flask and returned to the shaking incubator for the remainder of the experiment. At $t = 0$, the number of total viruses was quantified via plaque-forming units (PFUs) to determine the number of phages that contributed to subsequent infections. Sub-samples were taken every 20 min for 3 h and PFUs from free viruses (<0.2 µm filtrate) were quantified at each time point in duplicate using the agar-overlay method (Duhaime et al., 2011). The ends of the burst periods were determined where there was no significant change in the number of PFUs/ml from one time point to the next (Student's t -test). The burst size was calculated as the difference in PFUs/ml before the initial rise and after the first burst divided by the number of initially infecting phage.

Mass Spectroscopy-Based Structural Proteomics

Cesium chloride-purified phages were tryptically digested for 2D nano-LC-MS/MS analyses with an optimized Filter-Aided Sample Preparation kit (Expedeon, Inc., San Diego, CA, United States) (VerBerkmoes et al., 2009). MS/MS spectra were generated on a Velos Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, United States), as described (Holmfeldt et al., 2013). To recruit peptides to the phage genomes, spectra were searched using SEQUEST against a database consisting of the annotated phage proteins, all possible phage ORFs > 30 aa in all six reading frames, and eukaryotic organisms (human and mouse) to use as indicators for false positives. Data analyses were performed as described (Holmfeldt et al., 2013). A normalized spectral abundance factor (NSAF; Paoletti et al., 2006) was calculated for each structural protein of each phage (Supplementary Table S1).

Data from the mass spectroscopy-quantified peptide abundances was used to model the behavior of the phage structural proteins run on an SDS-PAGE gel. Band width in the model gel was based on the NSAF values for each protein with detected peptides. Band vertical position was based on molecular weights predicted from amino acid sequences. The data were compiled for visualization in a custom R script^{1,2}. For the largest band for each phage, 3-D predictions were modeled using I-TASSER (default settings) to test the hypothesis that they were phage major capsid proteins (Yang et al., 2015). The top-scoring model for each putative major capsid protein was further verified using ProSA to confirm that z-scores of all input structures were within the range typically found for native proteins of similar size (default settings; Wiederstein and Sippl, 2007). The Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB; Rose et al., 2017) was searched for the solved protein structure most similar to the structures predicted for the PSA phage major capsid proteins.

¹http://www-personal.umich.edu/~duhaimem/Rpubs_code/Helgoland_PSA_phage_Rpub.html

²<https://github.com/ecogenomix/PSAPhageGnms>

Sequencing and Annotation

Phage genomes were sequenced on the Illumina HiSeq platform and the PSA-HS4 genome closed using Sanger sequencing. ORFs were predicted using prodigal (default parameters; Hyatt et al., 2010). Annotations were made based on a combination of structural proteomics, domain identification in Pfam database (e-value < 0.001; Finn et al., 2010), BLASTP-identified (Altschul et al., 1990) homology to sequences in National Center for Biotechnology Information's (NCBI) non-redundant protein database (nr; January 2017), and tRNAs were searched for using tRNAscan-SE. Percentage of shared genes between the *Pseudoalteromonas* phages was based on an all-against-all comparison [BLASTP bit score > 75, per (Lavigne et al., 2008)]. Genome synteny plots were generated using Easyfig (v2.1; Sullivan et al., 2011) based on a full-genome BLASTN search, or occasionally TBLASTX where noted. The scope for whole-genome similarity to known phages was expanded beyond sequenced phage isolated to include prophages integrated in sequenced *Pseudoalteromonas* genomes. Prophages were predicted using VirSorter (version 1.0.2; Roux et al., 2015a) in RefSeq complete genomes and Whole Genome Sequencing (WGS) projects for organisms in the Order Alteromonadales, as well as sequenced Helgoland *Pseudoalteromonas* genomes (Duhaime et al., 2016). A heatmap was used to display relationships between *Pseudoalteromonas* phage and prophage genomes based on number of shared proteins. The orders of the genomes on the x- and y-axes were determined by the maximum distance method to encourage the self-hits and within-genus members to be attracted to the diagonal.

All predicted proteins from 43 global virus metagenomes, collectively the 'Tara Oceans Viromes (TOV)' dataset (Brum et al., 2015) were searched (blastp) against virus sequences in RefSeq (v70, May 2015). All TOV proteins with a bit score > 50 and e-value < 0.001 were considered already affiliated to existing known proteins. The remaining unaffiliated proteins were searched (blastp) against the 655 PSA phage proteins. Those with significant homology (bit > 50 and e-value < 1e-7) to PSA phage proteins were considered "newly affiliated." PCs that did not contain already-existing sequences from RefSeq were identified as "newly annotated" as PSA phage protein homologs. If one or more member was similar to a PSA phage protein, the PC was deemed newly affiliated.

All raw data and code for data analysis and figure generation have been made publically available by webserver³ and github repository⁴.

RESULTS AND DISCUSSION

Eight Helgoland *Pseudoalteromonas* Phages Constitute Five New Genera

The sequenced eight *Pseudoalteromonas* spp. phage genomes ranged in size from 35.3 to 129.4 kb and in G+C content from

35.7 to 44.7% (Table 1). Morphologically, the phages represented all three families in the order *Caudovirales*, with six siphoviruses, one podovirus, and one myovirus (Figure 1). These phages infect six closely related strains of *Pseudoalteromonas* sp. that varied by <2% 16S rRNA gene nucleotide identity (Figure 2A and Supplementary Figure S2).

The eight Helgoland *Pseudoalteromonas* phages belong to five new or candidate genera (delineated in Figure 1; described in Table 1) proposed here based on genome-based taxonomic guidelines (>40% proteins shared; Lavigne et al., 2008, 2009). Four of these Helgoland *Pseudoalteromonas* genera were novel, bringing the number of *Pseudoalteromonas* phage genera represented by isolated phages to a total of 11 (Figure 1). Seven of these are currently represented by a single phage. Only Helgoland podovirus PSA-HP1 formed a genus-level cluster with a previously sequenced phage isolate, *Pseudoalteromonas* phage RIO-1, which was isolated from the East Sea, South Korea, on 10 December 2007 (Hardies et al., 2013).

Phage genomes coincidentally sequenced during microbial host sequencing projects represent a recently illuminated source of phage genomic diversity (Roux et al., 2015b). As such, the 31 *Pseudoalteromonas* phages predicted to exist in sequenced *Pseudoalteromonas* genome projects (as prophages or extrachromosomal elements) were considered in this whole genome analysis (Figure 1; predicted prophages listed in Supplementary Table S2). When all *Pseudoalteromonas* phages and prophages are considered, 27 genera are resolved based on shared genome content, 18 of which are single phage or prophage genera (Supplementary Figure S1 and Table S1). The isolated *Pseudoalteromonas* phages and predicted *Pseudoalteromonas* prophages showed distinct clustering, which suggests minimal flow of genomic material between *Pseudoalteromonas* phages of these two contrasting lifestyles. Notably, the prophages include four from Helgoland *Pseudoalteromonas* hosts (strains H105, H103, and 13–15)—consistent with the dominant trend, none of these shared genome content with Helgoland phage isolates (Figure 1). The exceptions to this trend were Helgoland *Pseudomonas* phage PSA-HS6 and *Pseudomonas* phage PM2, two isolates that clustered in genera-level groups with predicted prophages (Figure 1). Both of these phages contain genes that suggest they may be capable of latent infections either through integration into their host genomes (PSA-HS6, discussed below) or as circularized plasmids (PM2; Männistö et al., 1999).

Of the 656 new *Pseudoalteromonas* phage proteins discovered in the Helgoland phage genomes, 285 were more similar to proteins in a global collection of ocean viral metagenomes—the combined 'POV' (Hurwitz and Sullivan, 2013) and 'TOV' (Brum et al., 2015) datasets—than to proteins in NCBI's RefSeq database. Of the sequences comprising existing PCs generated from POV and TOV (Hurwitz and Sullivan, 2013; Brum et al., 2015), 76 were homologous to the new *Pseudoalteromonas* phage proteins. This association informed 11 existing PCs that had been unannotated previously as bonafide phage clusters (Supplementary Table S3). Further, the PSA phage proteins were homologous to 422 singleton proteins that previously had not belonged to a virus PC, thereby forming new ocean virus PCs and shedding valuable light on a component of unknown virus sequence space.

³http://www-personal.umich.edu/~duhaimem/Rpubs_code/Helgoland_PSA_phage_Rpub.html

⁴<https://github.com/ecogenomix/PSAPhageGnms>

Identification of Phage Structural Proteomes

As structural proteins in new phages can rarely be annotated by sequence homology alone (Brum et al., 2016b), we chose to experimentally identify the proteins in purified viral particles using mass spectroscopy-based shotgun proteomics. One phage from each genus and two for the *PSAHS1*-likes were used to identify and annotate 94 proteins associated with the phage virions. Of these proteins, 49 were not previously known to be structural. Ten of the confirmed structural proteins belong to protein families with domains of unknown function (Pfam 'DUFs' 935, 1320, 4128, 4055, 3383, 2612; Supplementary Table S3), which help to propagate the structural annotations to a total of 840 DUF sequences in Pfam (counts as of January 2017). All but two of these DUF-containing structural proteins (*PSAHS1_00032* and *PSAHS4_00014*) affiliated with existing global ocean viral PCs (Supplementary Table S3). As they were found in the predominantly non-ocean phages and prophages

in public sequence databases as well as in ocean viromes, these data indicated that the core structural functions of PSA phages are conserved across diverse habitats.

The NSAF (Paoletti et al., 2006) was calculated for each proteome to quantitatively assess the relative abundances of the structural proteins comprising each phage. The proteins with the most abundant peptides in PSA-HP1, PSA-HM1, and PSA-HS6 were annotated as major capsid proteins or major capsid subunits based on their high peptide coverage (**Figure 3A**), as well as significant amino acid sequence homology with phage proteins in GenBank (Supplementary Table S3). Among the *PSAHS1*-like and *PSAHS2*-like siphoviruses, each representative proteome contained a single protein whose peptide coverage numerically dwarfed the others by 2- to 10-fold (**Figure 3A**). Based on the trend seen in the other PSA phages and the knowledge that major capsid proteins can constitute 57–65% of siphovirus proteomes (Buchwald et al., 1970; Zweig and Cummings, 1973) and are commonly the most abundant protein

TABLE 1 | Overview of Helgoland *Pseudoalteromonas* phage genera features.

| Family; proposed candidate genus | Phage | Genome size (bp) | G+C, % | ORF# | Phage isolation location | Genus features |
|----------------------------------|---|------------------|--------|------|------------------------------|--|
| Podoviridae; RIO1-like virus | PSA-HP1 | 45035 | 44.67 | 57 | North Sea | Members of the VpV262-like cluster have VpV262-like structural module arrangement (ter-port-scaff-cap) that differs from T7 organization (Hardies et al., 2013). Lineages of VpV262-likes have dissimilar replicative modules, see incongruence TerL (structural) and DNApol (replication) phylogenies (Supplementary Figures S4, S5). |
| | <i>Pseudoalteromonas</i> phage RIO-1 (Hardies et al., 2013) | 43882 | 44.7 | 56 | East Sea, South Korea | |
| Myoviridae; PSAHM1-like virus | PSA-HM1 | 129401 | 35.73 | 225 | North Sea | Shares 29% proteins and forms sub-family with PSA phage PH101 (Figure 6). Contains 27 T4-like genes (Supplementary Table S1); only one is structural. This is a novel non-T4-like myovirus genus. |
| Siphoviridae; PSAHS1-like virus | PSA-HS1 | 36769 | 40.52 | 62 | North Sea | Likely capable integrating into host genomes; contain ERF required for homologous recombination and Cro/C1 transcriptional repressor used to maintain lysogeny. |
| | PSA-HS4* (alias H105/1) | 38739 | 40.38 | 68 | | |
| | PSA-HS5 | 37230 | 40.39 | 62 | | |
| Siphoviridae; PSAHS2-like virus | PSA-HS2 | 37728 | 40.21 | 63 | North Sea | Lacks the canonical indicators of integration, e.g., in PSAHS1-likes |
| | PSA-HS8 | 37774 | 40.05 | 64 | | |
| Siphoviridae; PSAHS6-like virus | PSA-HS6 | 35328 | 40.21 | 54 | North Sea | PSA-HS6 is isolated; others are prophages. Hybrids of siphovirus non-contractile tail fused to a myo Mu-like head. Fusion also seen in <i>Pseudomonas</i> phages DS3112-likes (Figure 5B). Mu backbone likely ancestral (Wang et al., 2003); tail formation modules more recent. |
| | <i>Pseudoalteromonas</i> BSI20652, ANT505, and CP76 prophages | na | na | na | Spanish salterns, Antarctica | |

*PSA-HS4 is resequenced and now complete genome of *Pseudoalteromonas* phage H105/1 (RefSeq NC_015293).

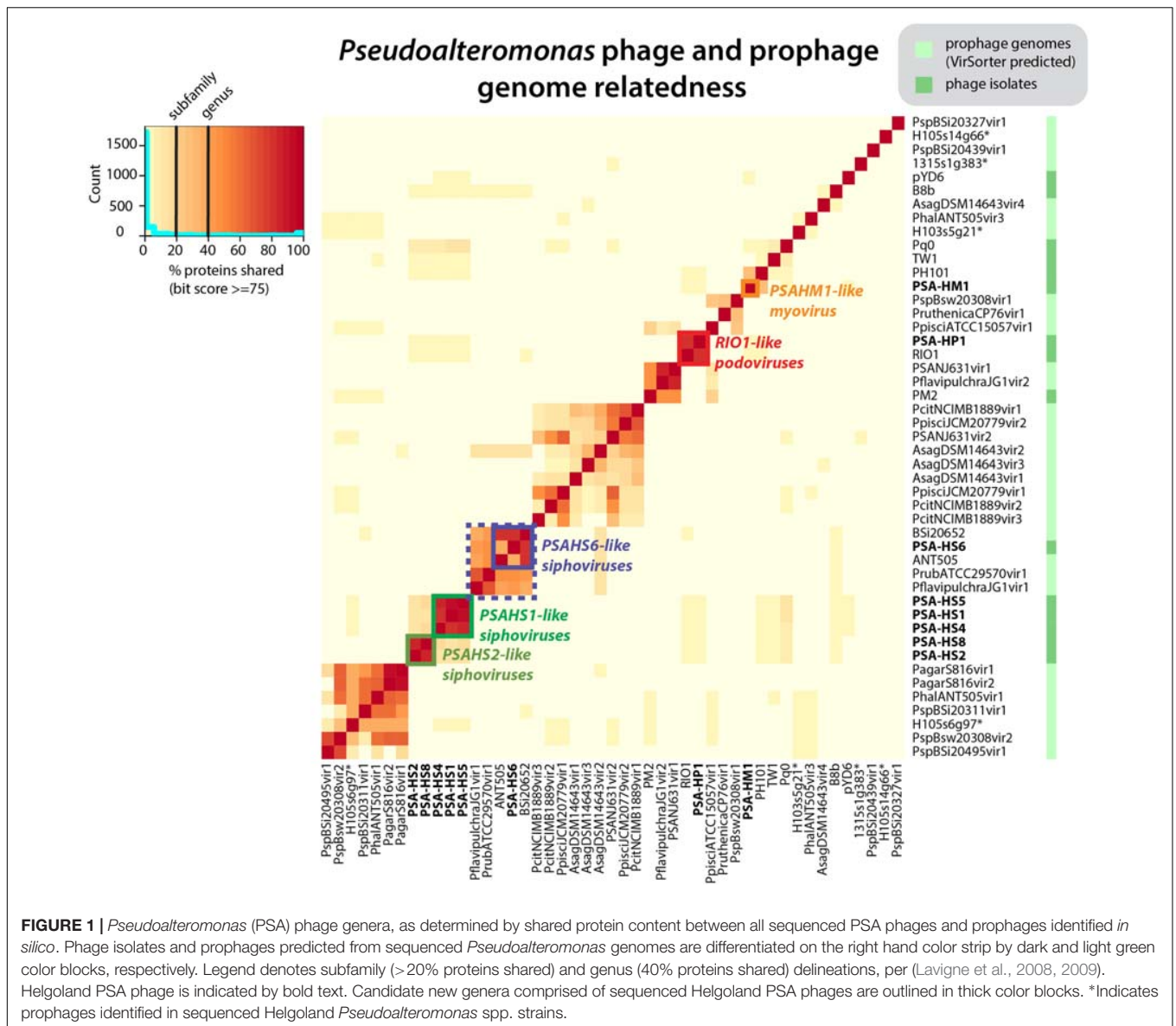
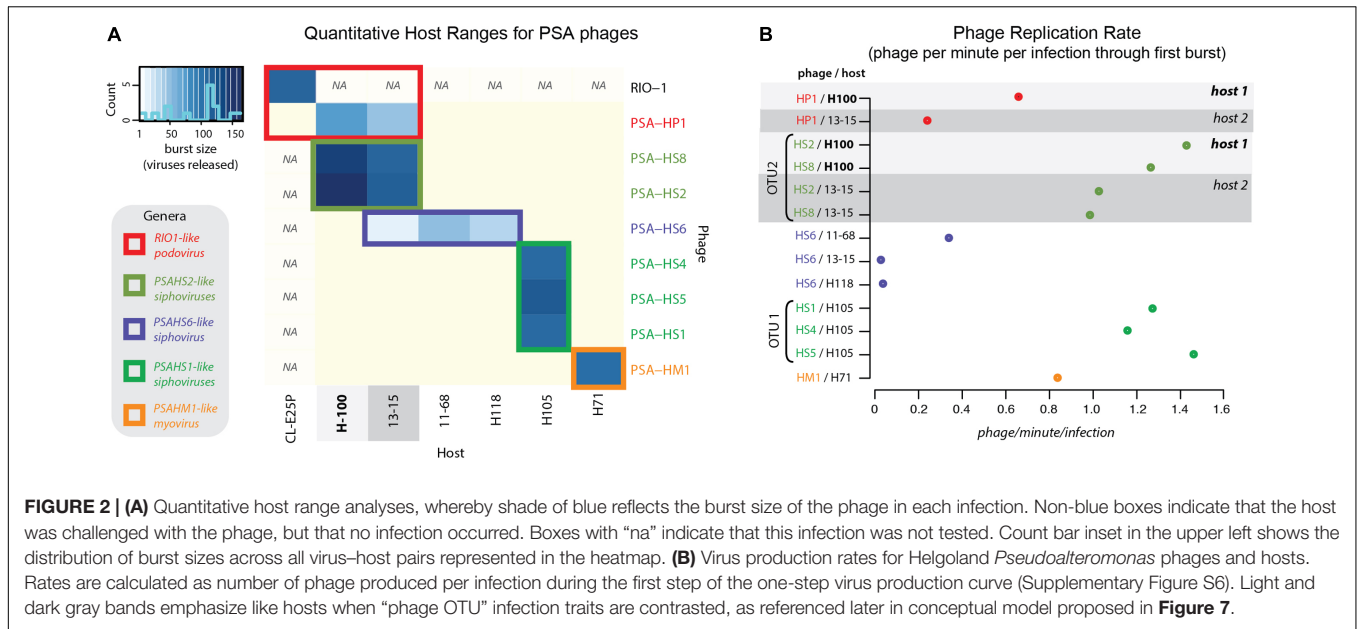


FIGURE 1 | *Pseudoalteromonas* (PSA) phage genera, as determined by shared protein content between all sequenced PSA phages and prophages identified *in silico*. Phage isolates and prophages predicted from sequenced *Pseudoalteromonas* genomes are differentiated on the right hand color strip by dark and light green color blocks, respectively. Legend denotes subfamily (>20% proteins shared) and genus (40% proteins shared) delineations, per (Lavigne et al., 2008, 2009). Helgoland PSA phage is indicated by bold text. Candidate new genera comprised of sequenced Helgoland PSA phages are outlined in thick color blocks. *Indicates prophages identified in sequenced Helgoland *Pseudoalteromonas* spp. strains.

of other phage families (Simoliūnas et al., 2013), we hypothesized that the high-abundance proteins and their homologs in the five *PSAHS1-like* and *PSAHS2-like* siphoviruses were major capsid proteins. This hypothesis was confirmed by the generation of 3D protein structure predictions of the putative PSA phage major capsid proteins followed by a comparison of these models with known structures in RCSB PDB (Rose et al., 2017). The z-scores of all predicted PSA major capsid protein structures were within the range of scores typically found for native proteins of similar size (Supplementary Figure S3; Wiederstein and Sippl, 2007), which provided confidence in the evidence based on structure prediction comparisons. For both the *PSAHS1-like* and *PSAHS2-like* Helgoland PSA siphoviruses, the best-aligned structure was a cryo-EM resolved protein, gp13, from *Bacillus subtilis* bacteriophage SPP1 (Figures 3B–D; White et al., 2012). The assembled capsid of phage SPP1 is comprised of multiple

gp13 hexamers that anchor a capsid spike protein, gp12 (White et al., 2012). Based on these results, we predict the capsid structure of the *PSAHS1-* and *PSAHS2-like* genera phages to be analogous to the SPP1 phage capsid (Figure 3D). These analyses enabled the annotation of previously unknown proteins as major capsid proteins in the five new Helgoland PSA siphoviruses.

Structural proteomics improved the resolution of structural module localization and, when combined with host range analysis, served to implicate three structural proteins as host range determinants. The *RIO1-like*s (PSA-HP1 and RIO-1), while highly syntenic, do not share five proteins in their structural modules that may be involved in host specificity (red outlined ORFs, Figure 4A). In our host range analyses, PSA-HP1 did not infect *Pseudoalteromonas marina* str. CL-E25P (Figure 2), the strain used to isolate RIO-1. We hypothesize that the proteins not shared between these phages may be involved in host



range determination. Contrastingly, podovirus PSA-HP1 and the *PSAHS2-like* siphoviruses have identical host ranges (**Figure 2**) and their only genotypic similarities are two structural proteins (PC49 and PC84, **Table 2**). Of these, one protein (PSA-HP1 ORF 17, PC49; black outlined ORFs, **Figure 4A**) is not found in the *PSAHS1-like* siphoviruses, which do *not* overlap in host range with PSA-HP1. We hypothesize that PC49 may play a role in host range and thereby bridging these *Pseudoalteromonas* phage families. We propose these candidates for future protein–protein interaction studies.

A Siphomyo Hybrid with Mu-Type Replication: A Rarely Described Feature of the Ocean Virus Landscape

The proposed *PSAHS6-like* genus contains siphovirus PSA-HS6 as its sole isolated member, but shares 90–92.4% ANI and 66–87% of its proteins with three prophages integrated in sequenced *Pseudoalteromonas* genomes in GenBank’s Whole Genome Shotgun (WGS) database (**Figures 1, 5A**). The *PSAHS6-like* phages share similarity with neither the *PSAHS1-like* nor *PSAHS2-like* Helgoland siphoviruses nor any other Helgoland phages (**Figure 1**). Rather, morphology and protein homology suggest the *PSAHS6-like* phages are modular hybrids composed of a siphovirus-like non-contractile tail fused to a myovirus Mu-like head (**Figure 5B**). Protein homology suggested the *PSAHS6-like* phages replicate by Mu-type transposition, which implies obligate integration into their host genomes and the capacity to remain as prophages in a host lysogen (Toussaint et al., 1994). The *PSAHS6-like* genomes contain numerous Mu-like conserved proteins localized in the two canonical Mu regions (Morgan et al., 2002): (i) head formation among the late genes and (ii) transposition and replication functions among the early genes (**Figure 5A**). Contrastingly, the *PSAHS6-like* tail formation module is homologous to non-Mu

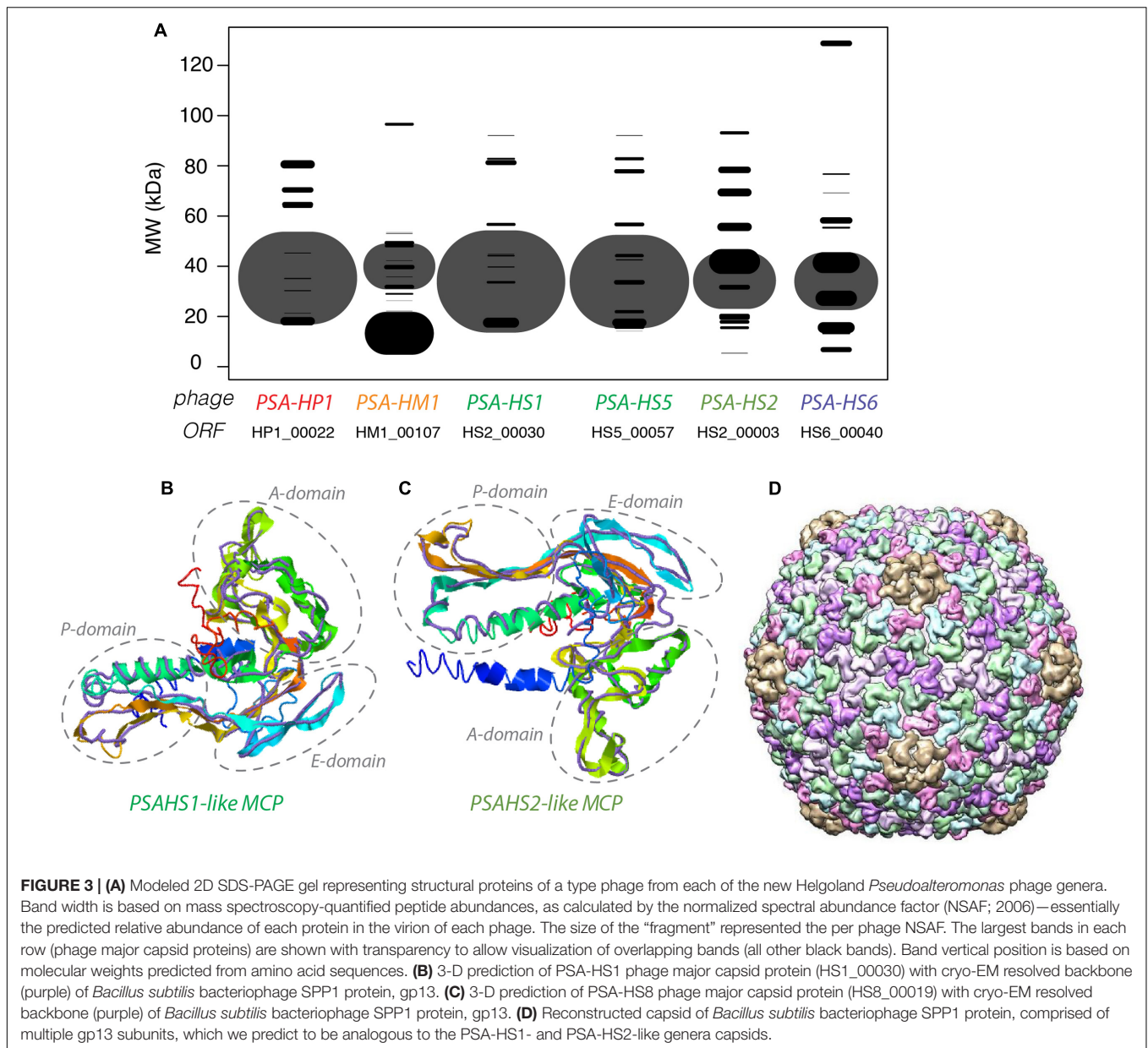
siphoviruses (e.g., *Shewanella* phage 1/44; **Figure 5B**). The similarity to Mu-like phages, supported by the location of the PSA-HS6 large terminase within the “Mu-like headful packaging” terminases (Supplementary Figure S5) and their propensity to be found integrated in host genomes, point to *PSAHS6-like* phages replicating via Mu-like transposition. Considering the propensity for Mu-like phages to package host DNA during transposition and replication (Toussaint et al., 1994) and that the *PSAHS6-like* superfamily is the second largest superfamily of *Pseudoalteromonas* phages (dotted box **Figure 1**), *PSAHS6-like* phages may be important components of the *Pseudoalteromonas* mobilome and sources of diversification and genomic exchange within this genus.

While similar Mu head-siphovirus tail fusions have been reported in a global collection of 12 phages and prophages of clinical *Pseudomonas aeruginosa* isolates (Wang et al., 2003; Cazares et al., 2014), descriptions of Mu-type siphomyo hybrids isolated from the marine environment are less common. Integrated Mu-like prophages are thought to be not inducible by mitomycin C (Paul, 2008)—though exceptions exist (Zheng et al., 2014). As such, quantifying their abundance in induced metagenome studies (e.g., McDaniel et al., 2008; Brum et al., 2016a) will be challenging. However, with phage and host in culture, we can now explore the environmental triggers that drive this Mu-type replicating virus to oscillate between lytic and lysogenic replication to better understand a class of lysogeny not yet explored in the environment.

Host Takeover Mechanisms Inferred from *Pseudoalteromonas* Phage Genomes

Hijacking Host Stress Response: MazG, DksA, PhoH

In order to respond to rapid environmental changes and resource availability, bacteria often adjust their global regulatory networks

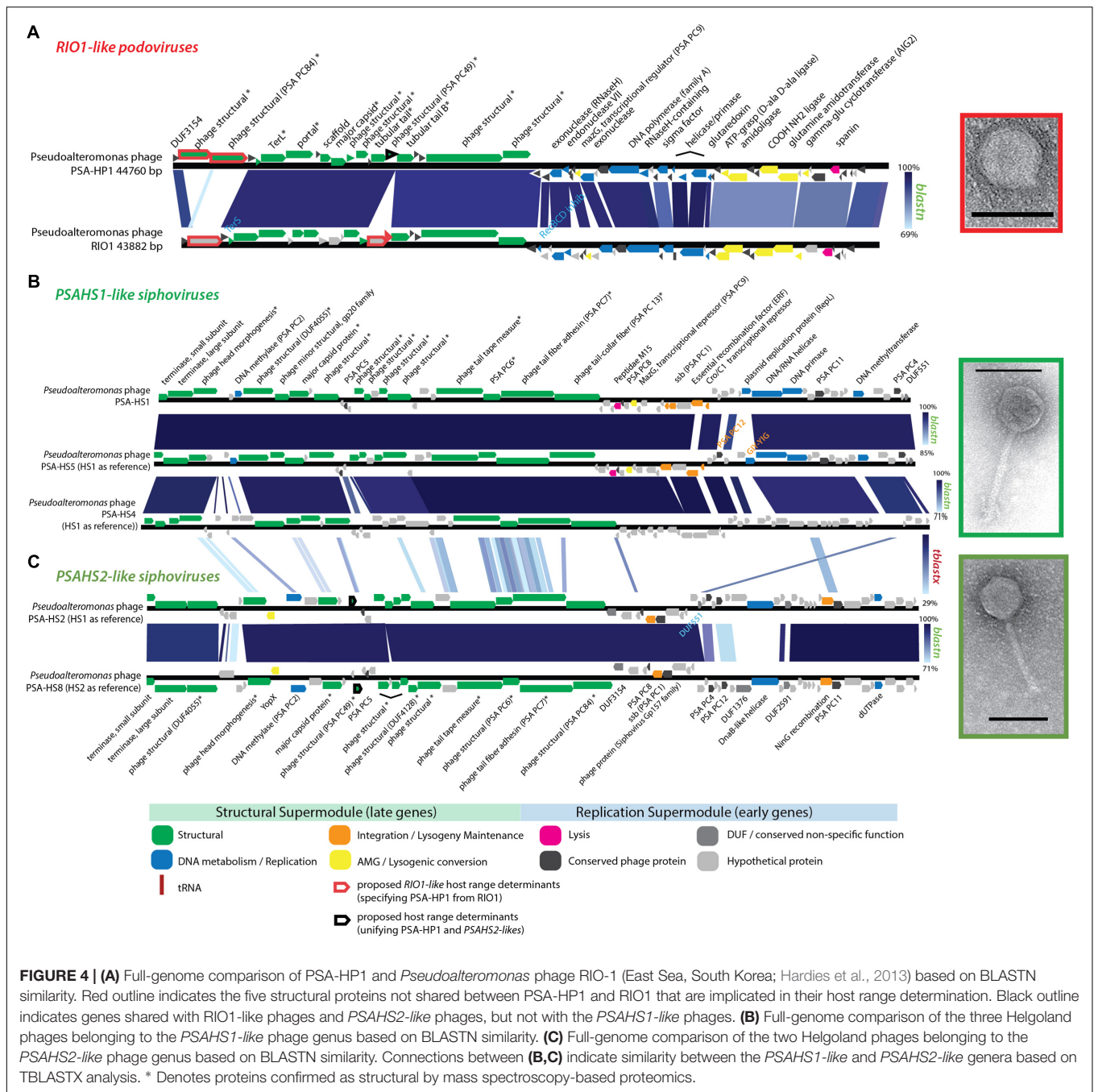


at the level of transcription (reviewed in Shimizu, 2013). One such mechanism is through the rapid production of a stringent response effector nucleotide, guanosine 3',5'-bispyrophosphate (ppGpp), in response to myriad nutritional stresses (Barker et al., 2001; Gross et al., 2006). ppGpp directly binds to promoter regions of RNA polymerase (RNAP) and globally influences transcription (Barker et al., 2001; Gross et al., 2006). Phages appear to have evolved mechanisms to tweak ppGpp-mediated response of host cells to the environment (Borbély et al., 1980; Clokie and Mann, 2006; Bryan et al., 2008).

MazG is a protein known to stall or reverse starvation-induced programmed cell death in *Escherichia coli* by decreasing the cellular pool of ppGpp (Gross et al., 2006). MazG, found in half the genomes presented here (podovirus PSA-HP1, the

three PSAHS1-like siphoviruses; **Figure 4B** and Supplementary Table S3), is over-represented in marine phages (Duhaime et al., 2011) and shared among all known T4-like cyanophages (Sullivan et al., 2010). When carried by phage and expressed during infection, MazG is thought to impede the global dampening of metabolic processes, thereby ensuring prolonged phage replication in a starving host (Clokie and Mann, 2006; Bryan et al., 2008). Further studies are needed to resolve the role of MazG in infection dynamics, particularly under conditions of nutrient-limitation induced host stress.

DksA is a critical component of the rRNA transcription initiation machinery that binds to RNA polymerase and also influences the regulation of rRNA promoters by ppGpp (Paul et al., 2004). The Helgoland *Pseudoalteromonas* Mu-like



siphovirus, PSA-HS6, and its relatives have DksA family transcriptional regulators (Figure 5A and Supplementary Table S3). We hypothesize that this system may provide a mechanism by which phage activity is modulated based on their host's physiological status, e.g., in the switch between lytic and lysogenic lifestyles.

PhoH is another protein commonly found in phage genomes. *phoH* is a core gene shared by 16 marine T4-like cyanophages compared in one study (Sullivan et al., 2010) and is found in other marine viruses (Goldsmith et al., 2011), including non-cyanobacterial myovirus Vibriophage KVP40 (Miller et al.,

2003), podovirus Roseophage SIO1 (Rohwer et al., 2000), and the Helgoland myovirus PSA-HM1 presented here (Figure 6 and Supplementary Table S3). This presence across diverse marine host types, a pattern that PSA-HM1 strengthens, hints toward *phoH* conferring a benefit to marine phages infecting ocean-dwelling hosts. PhoH belongs to the phosphate (*pho*) regulon, whose transcription is inducible under phosphate limitation (Kim et al., 1993). However, the specific dynamics of PhoH regulation are not uniform across systems (Goldsmith et al., 2011). For instance, during phosphate starvation the expression of the *phoH* gene increases in *Escherichia coli*

(Wanner, 1993) and *Corynebacterium glutamicum* (Ishige et al., 2003), decreases in marine *Synechococcus* (Tetu et al., 2009), and is unchanged in *Prochlorococcus* MED4 and 9313 (Martiny et al., 2006). In the absence of explicit phosphate-limitation in growth media, *phoH* is upregulated in *Prochlorococcus* MED4 during the late stages of infection by podovirus P-SSP7 (Lindell et al., 2007). We speculate that this may be due to cellular stress due to P-limitation as a result of the production of new virions that require more P (relative to C and N) than the ratio required by uninfected host cells (Jover et al., 2014). PhoH, believed to be a cytoplasmic protein involved in the uptake of phosphate under conditions of phosphate starvation (Kim et al., 1993; Makino et al., 1994), may confer an advantage to phages that carry it by supplementing phosphate uptake during the onset of P-limitation in the late stages of infection. Phosphorus indeed limits productivity in many ocean habitats and has been implicated as a dominant selective force in shaping microbial population heterogeneity (Coleman and Chisholm, 2010). Clearly, further experimental studies of infection under phosphate limitation are needed resolve these hypotheses and the role of phage-carried PhoH during infection.

Peptidoglycan Modification Cassette

The *RIO1-like* podoviruses carry a seven-gene cassette of peptidoglycan modification genes (PSA-HP1 ORFs 40–47; **Figure 4** and Supplementary Table S3) hypothesized to play a role in altering the host cell surface after infection to prevent superinfection by additional phages (Iyer et al., 2002). Similar genes have been found in other podoviruses (Hardies et al., 2013). Collectively these genes encode proteins

that perform functions necessary to synthesize three unusual linkages in peptidoglycan peptide side chains (Iyer et al., 2002; Hardies et al., 2013), including a gamma-glutamyl amidoligase (PSAHP1_00043), a second uncharacterized amidoligase (PSAHP1_00046), and an ATP grasp enzyme (PSAHP1_00047). Beyond synthesis, there are also peptidoglycan degradation genes encoded in the PSA-HP1 genome, including gamma-glutamyl cyclotransferase (PSAHP1_00040) and glutamine amidotransferase (PSAHP1_00042). This pathway is conserved across several marine and non-marine podovirus genomes, including *Pseudoalteromonas* phage RIO-1, Enterophage phiEco32, *Salmonella* phage 7-11, *Pseudomonas* phages PA11, tf, MR299-2, LUZ24, and PaP3, and *Salinivibrio* phage CW02 (Hardies et al., 2013). In these genomes, the host material recycling genes are located amidst genes involved in transcription, DNA metabolism, and replication – a pattern observed in other podoviruses (Lindell et al., 2007; Gao et al., 2012). Notably, genes in this region are co-transcribed in *Prochlorococcus* phage P-SSP7 replication (Lindell et al., 2007), suggesting that this shared genome organization could facilitate efficient recycling of limiting cellular resources during infection.

Advancing the Phage OTU Concept: Fitness-Determining Traits Are Conserved within Genome Groups of 95% ANI

An effective phage OTU definition should capture evolutionarily and ecologically cohesive populations (**Figure 7**). In other words, genotypes in a population at the defined OTU

TABLE 2 | Overview of proteins shared within and between genera.

| Family | Podo | | Myo | Sipho | | |
|--------------|------------------|--|-------------------------|--|-----------------|-----------------|
| | Genus | <i>RIO1-like</i> | <i>HM1-like</i> | <i>HS1-like</i> | <i>HS2-like</i> | <i>HS6-like</i> |
| Podo | <i>RIO1-like</i> | | | | | |
| Myo | <i>HM1-like</i> | None | | | | |
| Sipho | <i>HS1-like</i> | MazG nucleotide pyrophosphohydrolase (PSA PC9) | | | | |
| | <i>HS2-like</i> | Phage structural protein (PSA PC49) Phage structural protein (PSA PC84) | DNA methylase (PSA PC2) | Phage structural protein DNA methylase (PSA PC2) Phage head morphogenesis protein, SPP1 gp7 family Phage protein Phage protein (PSA PC11) Phage protein (PSA PC4) Phage tail fiber adhesin (PSA PC7) Phage structural protein (PSA PC6) Phage protein (PSA PC8) Phage structural protein (PSA PC84) Phage protein (PSA PC5) Phage tail tape measure protein | | |
| | <i>HS6-like</i> | None | None | None | None | |

Proteins shared between Helgoland *Pseudoalteromonas* phage groups with greater than or equal to 75 blast bit score or 40% amino acid identity, per (Lavigne et al., 2008, 2009).

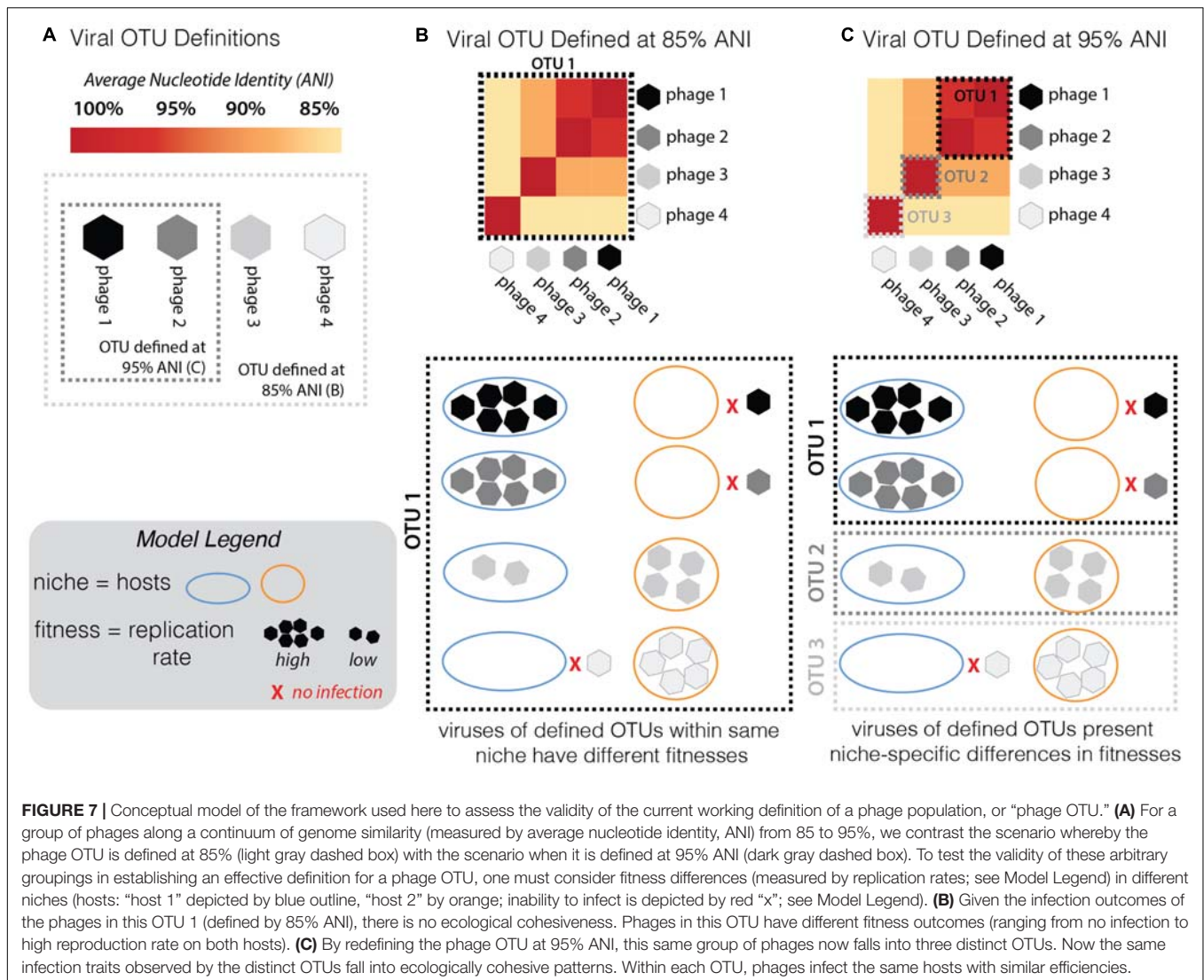


FIGURE 7 | Conceptual model of the framework used here to assess the validity of the current working definition of a phage population, or “phage OTU.” **(A)** For a group of phages along a continuum of genome similarity (measured by average nucleotide identity, ANI) from 85 to 95%, we contrast the scenario whereby the phage OTU is defined at 85% (light gray dashed box) with the scenario when it is defined at 95% ANI (dark gray dashed box). To test the validity of these arbitrary groupings in establishing an effective definition for a phage OTU, one must consider fitness differences (measured by replication rates; see Model Legend) in different niches (hosts: “host 1” depicted by blue outline, “host 2” by orange; inability to infect is depicted by red “x”; see Model Legend). **(B)** Given the infection outcomes of the phages in this OTU 1 (defined by 85% ANI), there is no ecological cohesiveness. Phages in this OTU have different fitness outcomes (ranging from no infection to high reproduction rate on both hosts). **(C)** By redefining the phage OTU at 95% ANI, this same group of phages now falls into three distinct OTUs. Now the same infection traits observed by the distinct OTUs fall into ecologically cohesive patterns. Within each OTU, phages infect the same hosts with similar efficiencies.

sensitive *Pseudoalteromonas* hosts were considered viral niches. Sensitivity of a host to a virus is not fully described by a binary relationship (e.g., infects/does not infect), but rather is a system-specific equation of, e.g., adsorption kinetics and infection mode (lytic/lysogenic decision), as demonstrated in marine *Cellulophaga* and its phages (Holmfeldt et al., 2014). Here, infection properties (burst sizes, latent periods, and replication rates) were tracked as traits indicative of viral fitness in the niche space tested (Figure 7). Notably, all infections were performed under identical media and culturing conditions. Theoretically, the full niche space of each virus would include the entire range of existing sensitive hosts, however, only part of this niche space could be explored here, as we were limited by the bacterial isolates available.

Previous studies have invoked a >95% ANI threshold to define and study the global biogeography of virus populations (Brum and Sullivan, 2015; Roux et al., 2016). In cyanophages, population genetic analyses support the validity of this cut-off

in delineating populations under different selective pressures and where barriers to cross-population gene flow exist (Gregory et al., 2016). Others have shown the existence of stable cyanophage ecotypes that persist for decades and that exhibit distinct temporal and spatial patterns of abundance—though in this study the clusters were defined at >93 and >98% ANI for their defined “clusters” and “subclusters,” respectively (Marston and Martiny, 2016). Yet, the appropriateness of the genome-similarity-based threshold remains unexplored in non-cyanophage systems and has not yet included consideration of fitness-determining infection traits. We hypothesized that if this OTU definition were meaningful in the *Pseudoalteromonas* phage–host system, trait-based differentiation would arise consistent with the 95% ANI genome-similarity boundaries used currently to define phage populations (Brum et al., 2015; Gregory et al., 2016).

Two multi-phage OTUs emerged from the Helgoland PSA phage collection. The siphoviruses in the *PSAHS1-like* genus (PSA-HS4 and PSA-HS5) share 97.6–99% genome-wide ANI

(**Figure 4B**) and those in the *PSAHS2-like* genus (PSA-HS2 and PSA-HS8) share 97.9% genome-wide ANI (**Figure 4C**). With this level of similarity, each set meet our criteria to be grouped into a single phage OTU (as in “OTU 1” in **Figure 7C**). The phenotypic trait data generated here for the multiple members of each phage OTU (e.g., burst sizes, host ranges, infection efficiencies; **Figure 2** and Supplementary Figure S2) suggest that the >95% ANI delineation represents meaningful ecological distinctions in this system. Specifically, when phages within each OTU have overlapping host ranges (i.e., occupy the same part of their niche), such as when HS2 and HS8 infect H100 (“host 1” in **Figure 2B**), they also have similar infection traits (burst sizes, latent periods, replication rates; **Figure 2**). However, this did not preclude different fitnesses when phages of an OTU occupied a different part of their niche, e.g., lower replication rates when phages of the *PSAHS2-like* OTU (PSA-HS2 and PSA-HS8) infected host 13–15 (“host 2”) instead of H100 (“host 1”; **Figure 2**). Phage fitness remained similar when phages of the same OTU occupied the same niche space (**Figure 2B**). These patterns provide evidence for niche-specific fitness-determining trait conservation within defined phage OTU populations.

On the contrary, while the PSA-HP1 and the observed niche space of the *PSAHS2-like* OTU population completely overlapped (i.e., infected the same hosts: H100 and 13–15; **Figure 2**), their fitnesses (i.e., phage replication rates) in this niche differ markedly (**Figure 7C**). In the niche space tested, these OTUs result in a distinct ecological impact. As the PSA-HP1 and the *PSAHS2-like* OTUs fail the genome-similarity criterion—as expected, considering their contrasting morphologies and thus structural gene sequences, they further represent “evolutionarily distinct” units.

While sample size is low and the full niche space of this set of OTUs has not been explored, these data and conceptual model suggest currently proposed genome-similarity thresholds delineate ecologically differentiated phage OTUs. Determination of fitness-conferring infection traits for the large collections of closely related phages (e.g., Pope et al., 2015; Gregory et al., 2016; Marston and Martiny, 2016) is a valuable metric that—while labor intensive to collect—would further test this OTU-defining model and enable the application of theory to advance viral community ecology.

CONCLUSION

The analysis of these eight new Helgoland *Pseudoalteromonas* phages is a first step in the development of a model system that will continue to improve our understanding of viruses infecting particle-associated ocean copiotrophs. Future work can build upon the comparative genomic foundation of this *Pseudoalteromonas* model system to investigate (i) the impact of boom–bust dynamics on the evolution and ecology of viruses of copiotrophs, including implications for population

genetics of both viruses and hosts, (ii) the adaptive mechanisms viruses have acquired to deal with their hosts’ responses to nutrient fluctuations in the environment, and (iii) the impacts of patchy infections in nature on both genome evolution and infection ecology. Beyond fundamental comparative genomic observations, exploration of genotype and phenotype linkages in the Helgoland *Pseudoalteromonas* phages offered insights into variability in infection traits within a genome-defined phage population, e.g., phage OTU. These data support a framework for future efforts of larger scale to evaluate the proposed phage OTU definition—a necessary building block in the pursuit to quantify the relative import of processes governing phage community ecology. Together with advances in virus–host interaction theory (Weitz et al., 2013), ecosystem models (Weitz et al., 2015), and approaches to integrate microbial omics and biogeochemical data at a global scale (Guidi et al., 2016), this new PSA phage–host model system provides insights into particle-associated virus–host interactions and brings us one step closer to developing a predictive understanding of how viruses alter natural ecosystems.

AUTHOR CONTRIBUTIONS

MD designed the experiments, performed analyses, and wrote the manuscript. NS performed laboratory experiments; acquired and analyzed infection data and critically edited the manuscript. SR performed bioinformatic analyses against POV and TOV and critically edited the manuscript. NV performed proteomic analyses. AW conceived experiments and critically edited the manuscript. MS designed experiments, critically edited the manuscript.

FUNDING

This work has been funded in part by University of Michigan ADVANCE Elizabeth Crosby award to MD, as well as Gordon and Betty Moore Foundation grants #2631, 3790, and 3305, and NSF OCE# 1536989 to MS.

ACKNOWLEDGMENT

We thank Dr. Byung Cheol Cho for providing the *Pseudoalteromonas marina* strain used to isolate phage RIO-1, as well as Jennifer Brum, Cristina Howard, and Vincent Deneff for critical review of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01241/full#supplementary-material>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Azam, F. (1998). Microbial control of oceanic carbon flux: the plot thickens. *Science* 280, 694–696. doi: 10.1126/science.280.5364.694
- Barker, M. M., Gaal, T., Josaitis, C. A., and Gourse, R. L. (2001). Mechanism of regulation of transcription initiation by ppppp. I. Effects of ppppp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* 305, 673–688. doi: 10.1006/jmbi.2000.4327
- Borbély, G., Kaki, C., Gulyás, A., and Farkas, G. L. (1980). Bacteriophage infection interferes with guanosine 3'-diphosphate-5'-diphosphate accumulation induced by energy and nitrogen starvation in the cyanobacterium *anacystis nidulans*. *J. Bacteriol.* 144, 859–864.
- Bratbak, G., Heldal, M., Norland, S., and Thingstad, T. F. (1990). Viruses as partners in spring bloom microbial trophodynamics. *Appl. Environ. Microbiol.* 56, 1400–1405.
- Breitbart, M. (2012). Marine viruses: truth or dare. *Annu. Rev. Mar. Sci.* 4, 425–448. doi: 10.1146/annurev-marine-120709-142805
- Brum, J. R., Hurwitz, B. L., Schofield, O., Ducklow, H. W., and Sullivan, M. B. (2016a). Seasonal time bombs: dominant temperate viruses affect southern ocean microbial dynamics. *ISME J.* 10, 437–449. doi: 10.1038/ismej.2015.125
- Brum, J. R., Ignacio-Espinoza, J. C., Kim, E. H., Trubl, G., Jones, R. M., Roux, S., et al. (2016b). Illuminating structural proteins in viral “dark matter” with metaproteomics. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2436–2441. doi: 10.1073/pnas.1525139113
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doucier, G., Acinas, S. G., Alberti, A., et al. (2015). Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498. doi: 10.1126/science.1261498
- Brum, J. R., and Sullivan, M. B. (2015). Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat. Rev. Microbiol.* 13, 147–159. doi: 10.1038/nrmicro3404
- Bryan, M. J., Burroughs, N. J., Spence, E. M., Clokie, M. R., Mann, N. H., and Bryan, S. J. (2008). Evidence for the intense exchange of MazG in marine cyanophages by horizontal gene transfer. *PLoS ONE* 3:e2048. doi: 10.1371/journal.pone.0002048
- Buchwald, M., Murialdo, H., and Siminovitch, L. (1970). The morphogenesis of bacteriophage lambda: II. Identification of the principal structural proteins. *Virology* 42, 390–400. doi: 10.1016/0042-6822(70)90282-5
- Cazaes, A., Mendoza-Hernández, G., and Guarneros, G. (2014). Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics* 15:1146. doi: 10.1186/1471-2164-15-1146
- Clokie, M. R., and Mann, N. H. (2006). Marine cyanophages and light. *Environ. Microbiol.* 8, 2074–2082. doi: 10.1111/j.1462-2920.2006.01171.x
- Coleman, M. L., and Chisholm, S. W. (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18634–18639. doi: 10.1073/pnas.1009480107
- Cordero, O. X., and Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* 12, 263–273. doi: 10.1038/nrmicro3218
- Deng, L., Ignacio-Espinoza, J. C., Gregory, A. C., Poulos, B. T., Weitz, J. S., Hugenholtz, P., et al. (2014). Viral tagging reveals discrete populations in *synechococcus* viral genome sequence space. *Nature* 513, 242–245. doi: 10.1038/nature13459
- Ducklow, H. W., Steinberg, D. K., and Buesseler, K. O. (2001). Upper ocean carbon export and the biological pump. *Oceanography* 14, 50–58. doi: 10.5670/oceanog.2001.06
- Duhaime, M. B., Wichels, A., and Sullivan, M. B. (2016). Six *Pseudoalteromonas* strains isolated from surface waters of Kabeltonne, offshore Helgoland, North Sea. *Genome Announc.* 4:e01697. doi: 10.1128/genomeA.01697-15
- Duhaime, M. B., Wichels, A., Waldmann, J., Teeling, H., and Glöckner, F. O. (2011). Ecogenomics and genome landscapes of marine *Pseudoalteromonas* phage H105/1. *ISME J.* 5, 107–121. doi: 10.1038/ismej.2010.94
- Falkowski, P. G., Fenchel, T., and DeLong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 320, 1034–1039. doi: 10.1126/science.1153213
- Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222. doi: 10.1093/nar/gkp985
- Fontanez, K. M., Eppley, J. M., Samo, T. J., Karl, D. M., and DeLong, E. F. (2015). Microbial community structure and function on sinking particles in the North Pacific subtropical gyre. *Front. Microbiol.* 6:469. doi: 10.3389/fmicb.2015.00469
- Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541–548. doi: 10.1038/21119
- Gao, E. B., Gui, J. F., and Zhang, Q. Y. (2012). A novel cyanophage with a cyanobacterial nonbleaching protein psbA gene in the genome. *J. Virol.* 86, 236–245. doi: 10.1128/JVI.06282-11
- Goldsmith, D. B., Crosti, G., Dwivedi, B., McDaniel, L. D., Varsani, A., Suttle, C. A., et al. (2011). Development of PhoH as a novel signature gene for assessing marine phage diversity. *Appl. Environ. Microbiol.* 77, 7730–7739. doi: 10.1128/AEM.05531-11
- Gregory, A. C., Solonenko, S. A., Ignacio-Espinoza, J. C., LaButti, K., Copeland, A., Sudek, S., et al. (2016). Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. *BMC Genomics* 17:930. doi: 10.1186/s12864-016-3286-x
- Gross, M., Marianovsky, I., and Glaser, G. (2006). MazG – a regulator of programmed cell death in *Escherichia coli*. *Mol. Microbiol.* 59, 590–601. doi: 10.1111/j.1365-2958.2005.04956.x
- Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature* 532, 465–470. doi: 10.1038/nature16942
- Hardies, S. C., Hwang, Y. J., Hwang, C. Y., Jang, G. I., and Cho, B. C. (2013). Morphology, physiological characteristics, and complete sequence of marine bacteriophage ΦRIO-1 infecting *Pseudoalteromonas marina*. *J. Virol.* 87, 9189–9198. doi: 10.1128/JVI.01521-13
- Holmfeldt, K., Howard-Varona, C., Solonenko, N., and Sullivan, M. B. (2014). Contrasting genomic patterns and infection strategies of two co-existing *Bacteroidetes* podovirus genera. *Environ. Microbiol.* 16, 2501–2513. doi: 10.1111/1462-2920.12391
- Holmfeldt, K., Solonenko, N., Shah, M., Corrier, K., Riemann, L., Verberkmoes, N. C., et al. (2013). Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12798–12803. doi: 10.1073/pnas.1305956110
- Hurwitz, B. L., Brum, J. R., and Sullivan, M. B. (2015). Depth-stratified functional and taxonomic niche specialization in the ‘core’ and ‘flexible’ Pacific Ocean Virome. *ISME J.* 9, 472–484. doi: 10.1038/ismej.2014.143
- Hurwitz, B. L., and Sullivan, M. B. (2013). The Pacific Ocean Virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS ONE* 8:e57355. doi: 10.1371/journal.pone.0057355
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Ishige, T., Krause, M., Bott, M., Wendisch, V. F., and Sahm, H. (2003). The phosphate starvation stimulon of *Corynebacterium glutamicum* determined by DNA microarray analyses. *J. Bacteriol.* 185, 4519–4529. doi: 10.1128/JB.185.15.4519-4529.2003
- Iyer, L. M., Koonin, E. V., and Aravind, L. (2002). Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSE transcription factors. *Genome Biol.* 3, 1–11. doi: 10.1186/gb-2002-3-3-research0012
- Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., and Weitz, J. S. (2014). The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat. Rev. Microbiol.* 12, 519–528. doi: 10.1038/nrmicro3289
- Kim, S. K., Makino, K., Amemura, M., Shinagawa, H., and Nakata, A. (1993). Molecular analysis of the PhoH gene, belonging to the phosphate regulon in *Escherichia coli*. *J. Bacteriol.* 175, 1316–1324. doi: 10.1128/jb.175.5.1316-1324.1993
- Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2567–2572. doi: 10.1073/pnas.0409727102

- Lavigne, R., Darius, P., Summer, E. J., Seto, D., Mahadevan, P., Nilsson, A. S., et al. (2009). Classification of myoviridae bacteriophages using protein sequence similarity. *BMC Microbiol.* 9:224. doi: 10.1186/1471-2180-9-224
- Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H. W., and Kropinski, A. M. (2008). Unifying classical and molecular taxonomic classification: analysis of the podoviridae using blastp-based tools. *Res. Microbiol.* 159, 406–414. doi: 10.1016/j.resmic.2008.03.005
- Lindell, D., Jaffe, J. D., Coleman, M. L., Futschik, M. E., Axmann, I. M., Rector, T., et al. (2007). Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* 449, 83–86. doi: 10.1038/nature06130
- Makino, K., Amemura, M., Kim, S. K., Nakata, A., and Shinagawa, H. (1994). Mechanism of transcriptional activation of the phosphate regulon in *Escherichia coli*. *Genes Dev.* 7, 149–160. doi: 10.1101/gad.7.1.149
- Männistö, R. H., Kivelä, H. M., Paulin, L., Bamford, D. H., and Bamford, J. K. (1999). The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. *Virology* 262, 355–363. doi: 10.1006/viro.1999.9837
- Marston, M. F., and Amrich, C. G. (2009). Recombination and microdiversity in coastal marine cyanophages. *Environ. Microbiol.* 11, 2893–2903. doi: 10.1111/j.1462-2920.2009.02037.x
- Marston, M. F., and Martiny, J. B. (2016). Genomic diversification of marine cyanophages into stable ecotypes. *Environ. Microbiol.* 18, 4240–4253. doi: 10.1111/1462-2920.13556
- Martiny, A. C., Coleman, M. L., and Chisholm, S. W. (2006). Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12552–12557. doi: 10.1073/pnas.0601301103
- McDaniel, L., Breitbart, M., Mobberley, J., Long, A., Haynes, M., Rohwer, F., et al. (2008). Metagenomic analysis of lysogeny in Tampa Bay: implications for prophage gene expression. *PLoS ONE* 3:e3263. doi: 10.1371/journal.pone.0003263
- Miller, E. S., Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Durkin, A. S., Ciecko, A., et al. (2003). Complete genome sequence of the broad-host-range Vibriophage KVP40: Comparative genomics of a T4-related bacteriophage. *J. Bacteriol.* 185, 5220–5233. doi: 10.1128/JB.185.17.5220-5233.2003
- Moebus, K. (1992). Further investigations on the concentration of marine bacteriophages in the water around Helgoland, with E to the phage-host systems encountered. *Helgoländer Meeresunters.* 46, 275–292. doi: 10.1007/BF02367099
- Morgan, G. J., Hatfull, G. F., Casjens, S., and Hendrix, R. W. (2002). Bacteriophage mu genome sequence: analysis and comparison with mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. *J. Mol. Biol.* 317, 337–359. doi: 10.1006/jmbi.2002.5437
- Paoletti, A. C., Parmely, T. J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R. C., et al. (2006). Quantitative proteomic analysis of distinct mammalian mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18928–18933. doi: 10.1073/pnas.0606379103
- Paul, B. J., Barker, M. M., Ross, W., Schneider, D. A., Webb, C., Foster, J. W., et al. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* 118, 311–322. doi: 10.1016/j.cell.2004.07.009
- Paul, J. H. (2008). Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2, 579–589. doi: 10.1038/ismej.2008.35
- Ploug, H., Grossart, H. P., Azam, F., and Jorgensen, B. B. (1999). Photosynthesis, respiration, and carbon turnover in sinking marine snow from surface waters of Southern California Bight: implications for the carbon cycle in the ocean. *Mar. Ecol. Prog. Ser.* 179, 1–11. doi: 10.3354/meps179001
- Polz, M. F., Hunt, D. E., Preheim, S. P., and Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361, 2009–2021. doi: 10.1098/rstb.2006.1928
- Pope, W. H., Bowman, C. A., Russell, D. A., Jacobs-Sera, D., Asai, D. J., Cresawn, S. G., et al. (2015). Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *ELife.* 4:e06416. doi: 10.7554/eLife.06416
- Proctor, L. M., and Fuhrman, J. A. (1991). Roles of viral infection in organic particle flux. *Mar. Ecol. Prog. Ser.* 69, 133–142
- Rohwer, F., Segall, A., Steward, G., Seguritan, V., Breitbart, M., Wolven, F., et al. (2000). The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with nonmarine phages. *Limnol. Oceanogr.* 45, 408–418. doi: 10.4319/lo.2000.45.2.0408
- Rose, P. W., Prlič, A., Altunkaya, A., Bi, C., Bradley, A. R., Christie, C. H., et al. (2017). The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* 45, D271–D281. doi: 10.1093/nar/gkw1000
- Roux, S., Brum, J. R., Dutilh, B. E., Sunagawa, S., Duhaime, M. B., Loy, A., et al. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537, 689–693. doi: 10.1038/nature19366
- Roux, S., Enault, F., Hurwitz, B. L., and Sullivan, M. B. (2015a). VirSorter: mining viral signal from microbial genomic data. *PeerJ.* 3:e985. doi: 10.7717/peerj.985
- Roux, S., Hallam, S. J., Woyke, T., and Sullivan, M. B. (2015b). Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *ELife.* 4:e08490. doi: 10.7554/eLife.08490
- Shanks, A. L., and Trent, J. D. (1980). Marine snow: sinking rates and potential role in vertical flux. *Deep Sea Res.* 27, 137–143. doi: 10.1016/0198-0149(80)90092-8
- Shimizu, K. (2013). Regulation systems of bacteria such as *Escherichia coli* in response to nutrient limitation and environmental stresses. *Metabolites* 4, 1–35. doi: 10.3390/metabo4010001
- Simoliūnas, E., Kaliniene, L., Truncaitė, L., Zajanėkauskaitė, A., Staniulis, J., Kaupinis, A., et al. (2013). Klebsiella phage vb_klem-rak2 - a giant singleton virus of the family myoviridae. *PLoS ONE* 8:e60717. doi: 10.1371/journal.pone.0060717
- Smrga, S., Fernandez, V. I., Mitchell, J. G., and Stocker, R. (2016). Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 113, 1576–1581. doi: 10.1073/pnas.1512307113
- Stocker, R., Seymour, J. R., Samadani, A., Hunt, D. E., and Polz, M. F. (2008). Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4209–4214. doi: 10.1073/pnas.0709765105
- Sullivan, M. B., Huang, K. H., Ignacio-Espinoza, J. C., Berlin, A. M., Kelly, L., Weigle, P. R., et al. (2010). Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ. Microbiol.* 12, 3035–3056. doi: 10.1111/j.1462-2920.2010.02280.x
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750
- Tetu, S. G., Brahamsha, B., Johnson, D. A., Tai, V., Phillippy, K., Palenik, B., et al. (2009). Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. *ISME J.* 3, 835–849. doi: 10.1038/ismej.2009.31
- Thomas, T., Evans, F. F., Schleheck, D., Mai-Prochnow, A., Burke, C., Penesyan, A., et al. (2008). Analysis of the *Pseudoalteromonas tunicata* genome reveals properties of a surface-associated life style in the marine environment. *PLoS ONE* 3:e3252. doi: 10.1371/journal.pone.0003252
- Toussaint, A., Gama, M.-J., Laachouch, J., Maenhaut-Michel, G., and Mhammedi-Alaoui, A. (1994). Regulation of bacteriophage Mu transposition. *Genetica* 93, 27–39. doi: 10.1007/BF01435237
- VerBerkmoes, N. C., Denev, V. J., Hettich, R. L., and Banfield, J. F. (2009). Systems biology: functional analysis of natural microbial consortia using community proteomics. *Nat. Rev. Microbiol.* 7, 196–205. doi: 10.1038/nrmicro2080
- Wang, P. W., Chu, L., and Guttman, D. S. (2003). Complete sequence and evolutionary genomic analysis of the *Pseudomonas aeruginosa* transposable bacteriophage D3112. *J. Bacteriol.* 186, 400–410. doi: 10.1128/JB.186.2.400-410.2004
- Wanner, B. L. (1993). Gene regulation by phosphate in enteric bacteria. *J. Cell. Biochem.* 51, 47–54. doi: 10.1002/jcb.240510110
- Weinbauer, M. G., Bettarel, Y., Cattaneo, R., Luef, B., Maier, C., Motegi, C., et al. (2009). Viral ecology of organic and inorganic particles in aquatic systems: avenues for further research. *Aquat. Microb. Ecol.* 57, 321. doi: 10.3354/ame01363

- Weitz, J. S., Poisot, T., Meyer, J. R., Flores, C. O., Valverde, S., Sullivan, M. B., et al. (2013). Phage-bacteria infection networks. *Trends Microbiol.* 21, 82–91. doi: 10.1016/j.tim.2012.11.003
- Weitz, J. S., Stock, C. A., Wilhelm, S. W., Bourouiba, L., Coleman, M. L., Buchan, A., et al. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J.* 9, 1352–1364. doi: 10.1038/ismej.2014.220
- White, H. E., Sherman, M. B., Brasilès, S., Jacquet, E., Seavers, P., Tavares, P., et al. (2012). Capsid structure and its stability at the late stages of bacteriophage SPP1 assembly. *J. Virol.* 86, 6768–6777. doi: 10.1128/JVI.00412-12
- Wichels, A., Biel, S. S., Gelderblom, H. R., Brinkhoff, T., Muyzer, G., and Schütt, C. (1998). Bacteriophage diversity in the North Sea. *Appl. Environ. Microbiol.* 64, 4128–4133.
- Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, 407–410. doi: 10.1093/nar/gkm290
- Wommack, K. E., and Colwell, R. R. (2000). Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213
- Yu, Z. C., Chen, X. L., Shen, Q. T., Zhao, D. L., Tang, B. L., Su, H. N., et al. (2015). Filamentous phages prevalent in *Pseudoalteromonas* spp. confer properties advantageous to host survival in arctic sea ice. *ISME J.* 9, 871–881. doi: 10.1038/ismej.2014.185
- Zheng, Q., Zhang, R., Xu, Y., White, R. A., Wang, Y., Luo, T., et al. (2014). A marine inducible prophage vB_CibM-P1 isolated from the aerobic anoxygenic phototrophic bacterium *Citromicrobium bathyomarinum* JL354. *Sci. Rep.* 4:7118. doi: 10.1038/srep07118
- Zweig, M., and Cummings, D. J. (1973). Structural proteins of bacteriophage T5. *Virology* 51, 443–453. doi: 10.1016/0042-6822(73)90443-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Duhaime, Solonenko, Roux, Verberkmoes, Wichels and Sullivan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.