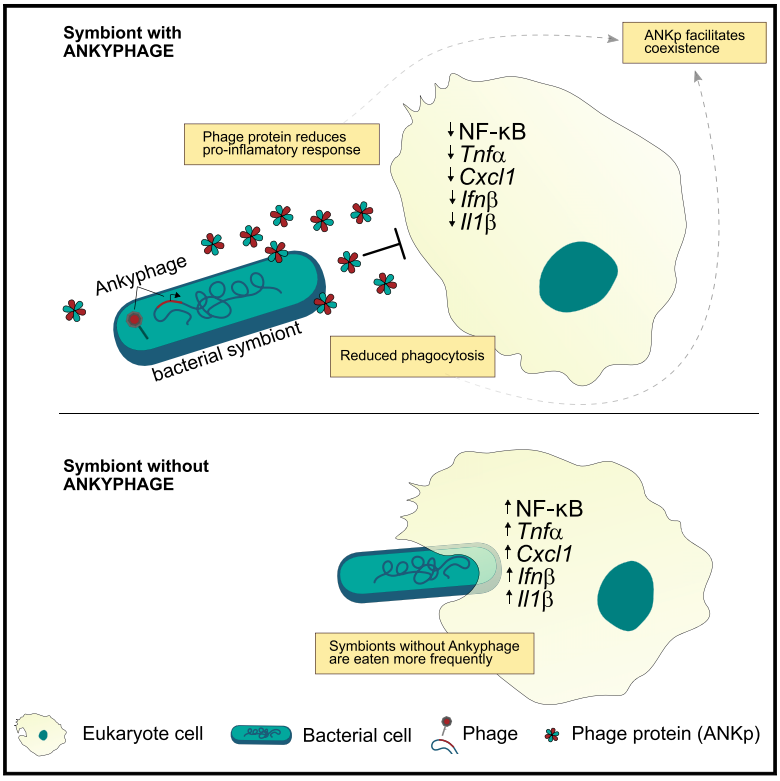
Short Article

Cell Host & Microbe

A Phage Protein Aids Bacterial Symbionts in Eukaryote Immune Evasion

Graphical Abstract



Highlights

- Sponges, evolutionary basal animals, represent a reservoir of novel viral diversity
- Viromes of neighboring sponges are individually unique and species specific
- Phages encode ankyrins to aid bacteria in evading the eukaryotic immune system
- Such "Ankyphages" are widespread in host-associated environments, including humans

Authors

Martin T. Jahn, Ksenia Arkhipova, Sebastian M. Markert, ..., Philip Rosenstiel, Bas E. Dutilh, Ute Hentschel

Correspondence

mjahn@geomar.de (M.T.J.), uhentschel@geomar.de (U.H.)

In Brief

Jahn et al. find that sponges, although massively filtering seawater, host individually unique and species-specific viral communities. An abundant sponge bacteriophage encodes ankyrins that, upon bacterial expression, reduce the eukaryotic immune response and phagocytosis of bacteria. This suggests a mechanism of tripartite phage-bacterium-host interplay with the phage fostering host-microbe symbiosis.

Jahn et al., 2019, Cell Host & Microbe *26*, 1–9
October 9, 2019
https://doi.org/10.1016/j.chom.2019.08.019

CellPress

Please cite this article in press as: Jahn et al., A Phage Protein Aids Bacterial Symbionts in Eukaryote Immune Evasion, Cell Host & Microbe (2019), https://doi.org/10.1016/j.chom.2019.08.019

Cell Host & Microbe
Short Article

A Phage Protein Aids Bacterial Symbionts in Eukaryote Immune Evasion

Martin T. Jahn,[1,*] Ksenia Arkhipova,[2] Sebastian M. Markert,[3] Christian Stigloher,[3] Tim Lachnit,[4] Lucia Pita,[1] Anne Kupczok,[4] Marta Ribes,[5] Stephanie T. Stengel,[6] Philip Rosenstiel,[4,6] Bas E. Dutilh,[2] and Ute Hentschel[1,4,7,*]
[1]GEOMAR Helmholtz Centre for Ocean Research Kiel, Marine Symbioses, 24105 Kiel, Germany
[2]Theoretical Biology and Bioinformatics, Utrecht University, 3584 Utrecht, the Netherlands
[3]Imaging Core Facility, Biocenter, University of Würzburg, 97074 Würzburg, Germany
[4]Christian-Albrechts-University of Kiel, 24105 Kiel, Germany
[5]Institut de Ciències del Mar-CSIC, 08003 Barcelona, Spain
[6]Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, 24105 Kiel, Germany
[7]Lead Contact
*Correspondence: mjahn@geomar.de (M.T.J.), uhentschel@geomar.de (U.H.)
https://doi.org/10.1016/j.chom.2019.08.019

SUMMARY

Phages are increasingly recognized as important members of host-associated microbiomes, with a vast genomic diversity. The new frontier is to understand how phages may affect higher order processes, such as in the context of host-microbe interactions. Here, we use marine sponges as a model to investigate the interplay between phages, bacterial symbionts, and eukaryotic hosts. Using viral metagenomics, we find that sponges, although massively filtering seawater, harbor species-specific and even individually unique viral signatures that are taxonomically distinct from other environments. We further discover a symbiont phage-encoded ankyrin-domain-containing protein, which is widely spread in phages of many host-associated contexts including human. We confirm in macrophage infection assays that the ankyrin protein (ANKp) modulates the eukaryotic host immune response against bacteria. We predict that the role of ANKp in nature is to facilitate coexistence in the tripartite interplay between phages, symbionts, and sponges and possibly many other host-microbe associations.

INTRODUCTION

Phages are the most abundant and diverse entities in the oceans (Gregory et al., 2019; Rohwer, 2003; Wommack and Colwell, 2000) and, along with their role as major bacterial killers, significantly impact global biochemical cycles (Breitbart et al., 2018; Suttle, 2007), bacterial fitness, and diversity (Betts et al., 2018; Marston et al., 2012). A plethora of fine-tuned defense and counter-defense mechanisms and lysogenic conversion factors have been discovered through research focusing on phage-bacteria interactions (Barrangou et al., 2007; Kronheim et al., 2018). Importantly, however, in host-associated microbial communities, a third player, the eukaryotic host, not only sets the stage but may also interact with both other parties in its own interest. Surprisingly little is, however, known about the tripartite interaction between phages, their bacterial hosts, and the animals that harbor the microbial communities (Barr et al., 2013; Keen and Dantas, 2018).

Marine sponges and their dense and diverse microbial symbiont communities are attractive models for the study of host-microbe-phage interactions (Thomas et al., 2016). As filter-feeding animals, sponges pump up to 24,000 liters of seawater through their system per day (Weisz et al., 2008), exposing them to up to an estimated $\sim 2.4 \times 10^{13}$ viruses daily. Interestingly, defense mechanisms (e.g., restriction modification and CRISPR-Cas) to selfish genetic elements (e.g., phages and plasmids) are clearly enriched in sponge symbiont genomes (Horn et al., 2016; Podell et al., 2019; Slaby et al., 2017).

While these features indicate phage resistance of bacterial sponge symbionts, the beneficial effects of phages on the sponge microbial community are largely unexplored. The first and only other sponge virome sequencing approach revealed species-specific viral signatures in Great Barrier Reef sponges, which shared low identity to known viral genomes (Laffy et al., 2018). Here, we generated nested viromes from Mediterranean sponges and identified specialized phage taxa and host-enriched phage functions. Importantly, we discovered and expressed an immunomodulatory phage protein that critically alters microbe-eukaryote interactions with potential implications in sponges and many other systems.

RESULTS

High Diversity and Novelty in Marine Sponge Viromes

We report the metagenomic analysis of marine sponge viromes that were sampled to cover the levels of sponge species, sponge individuals, and sponge tissues (outer layer "pinacoderm" and inner tissue "mesohyl matrix"). Viruses from nearby seawater, collected in immediate vicinity and at the same time, were used as controls. With 142 Gbp of sequencing data from 32 sponges (two tissues × 4 individuals × 4 species) and 4 seawater reference viromes, this represents the deepest sequencing effort performed on sponge viruses to date. The final assembly contained 4,484 curated viral contigs representing population-level genomes (≥5 kb, hereafter termed "BCvir"

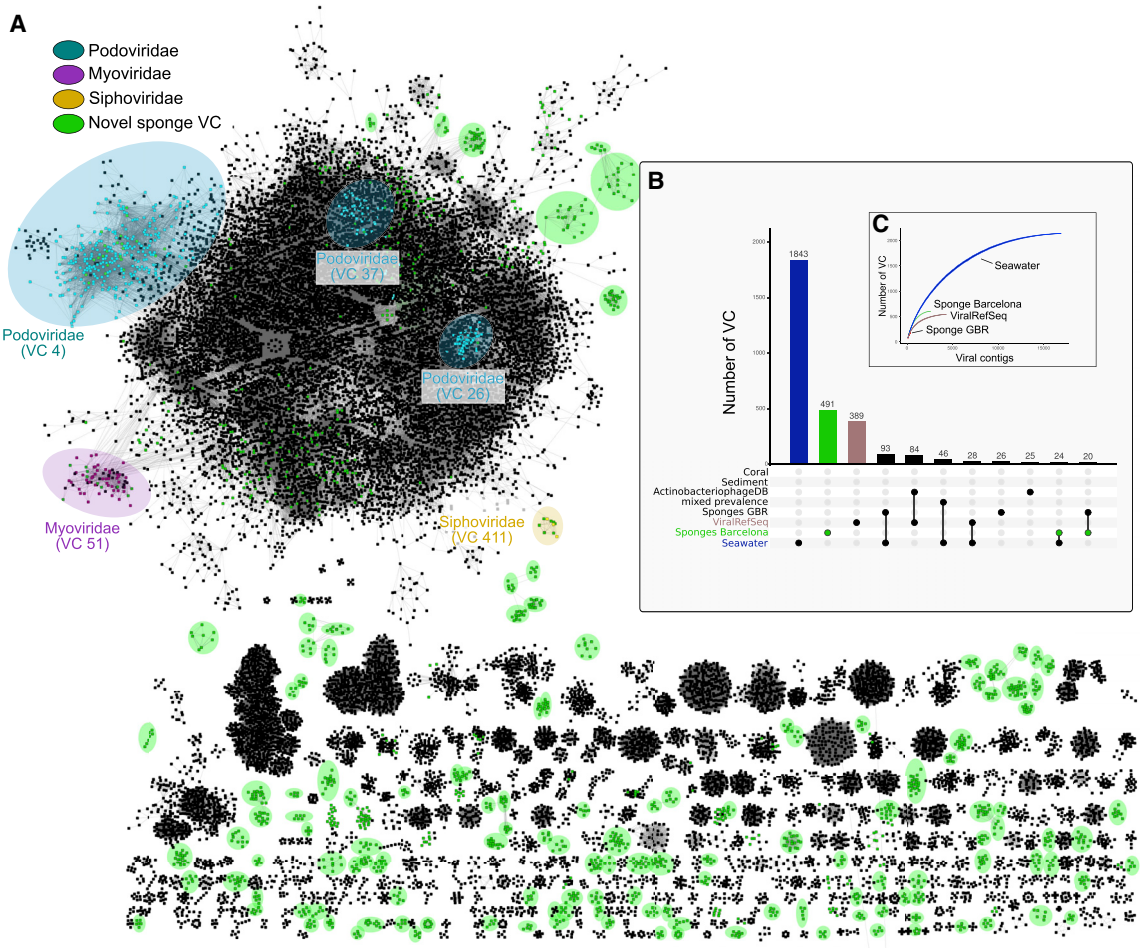


Figure 1. Diversity of Sponge Viruses in Relation to Known Viral Sequence Space

(A) The gene-sharing network associates identified sponge viral clusters in this study (VCs, green ovals) with a database of viral genomes that includes (i) assembled contigs from 130 marine viromes, (ii) the ActinophageDB, and (iii) ViralRefseq (black). Further, VCs containing the 5 most taxonomically assigned sponge viral genomes are highlighted. Network nodes are viral genomes and edges are significant similarities between them based on shared gene content. VCs with at least 5 genomes are shown (see detailed network statistics in Table S2 and raw data in Data S1).

(B) The matrix layout shows the number of VCs that are exclusive (one circle) or shared (multiple circles) between the eight different datasets used for clustering. Shown are the top intersections ($\geq$20 members) as a vertical bar plot, sorted by size.

(C) Rarefaction curves for the most diverse datasets showing the accumulation of VCs as a function of sampled viral contigs (N).

for viral populations of the North Western Mediterranean Coast close to Barcelona), representing 51.4% of all the read-level data (Table S1; Figure S1). Of these, 101 were circular with matching ends and represent putatively complete viral genomes. The remaining contigs (of which 1,649 were $\geq$10 kb) were either putative linear genomes or genome fragments (Roux et al., 2019). To investigate how the 4,484 BCvir populations were positioned in the known viral sequence space, we clustered our sequences with an extended sequence space of 11,901 viral genome sequences obtained from viral Refseq and the ActinophageDB as well as 29,922 assembled contigs from 130 publicly available marine viral communities. This analysis was based on shared gene content and detected 3,218 viral clusters (VCs) (Figure 1A). The 4,484 BCvir populations partitioned into 813 VCs (green) representing 25.3% (n = 813 of 3,218 VCs) of the total viral diversity included in this extended database. Notably, most of the BCvir diversity consisted of viruses that were never detected before, as indicated by the fact that these VCs contained only BCvir contigs (n = 491 of 3,218 VCs; 15.3%) (Figure 1B), many of which shared no distant edge with other VCs. To ensure that this observation was not inflated by the shorter 5 kb contig length cutoff, which we had initially applied to capture shorter single-stranded DNA (ssDNA) viruses, we performed the same analysis again using a more stringent 10 kb length filter as suggested in Roux et al. (2017). With this approach, the 1,649 BCvir populations partitioned into 997 sponge VCs, of which 371 were uncharacterized (n = 371 of 1,304 VCs; 28.5%) in the extended sequence space. VCs delineate approximately genus-level taxonomy in known viruses (Lima-Mendez et al., 2008; Roux et al., 2015) with at least 371 sponge-derived VCs appear not to be part of the 803 viral genera currently listed by the International Committee on Taxonomy of Viruses ICTV (King et al., 2018) (via ViralRefseq; see Figure S2). Our virome dataset contained 3.9% BCvir

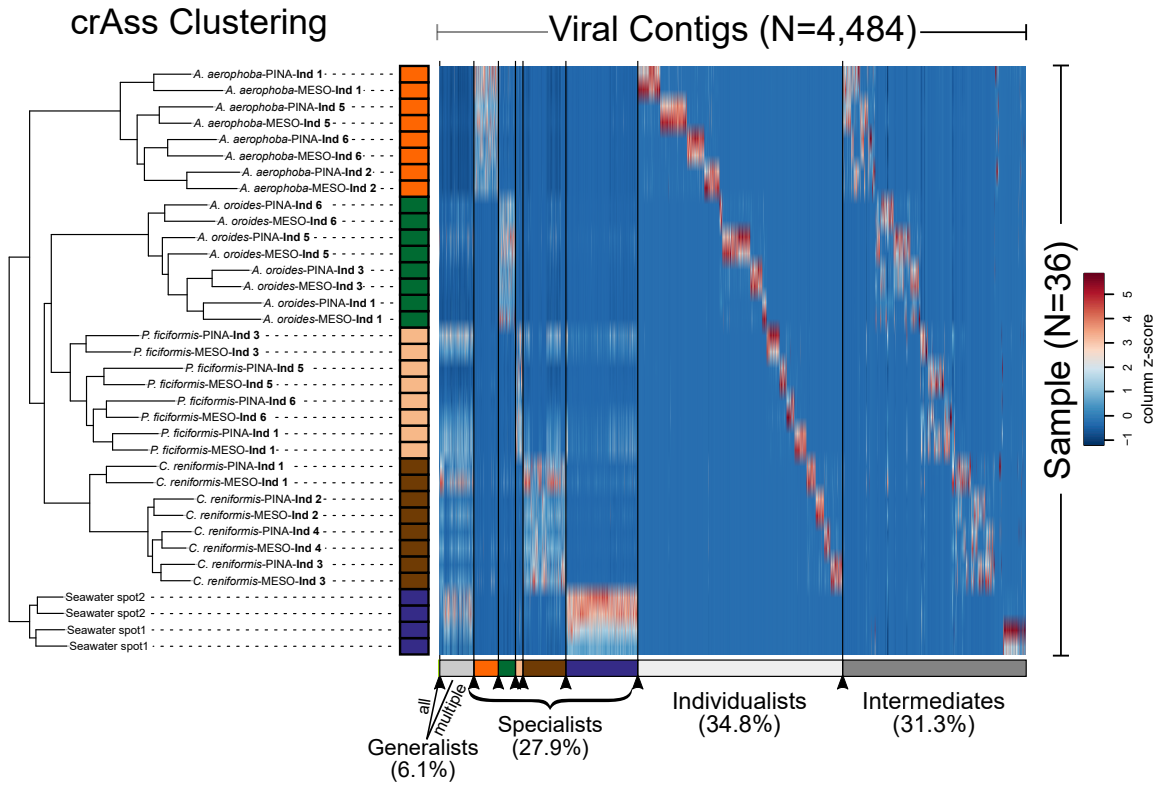


Figure 2. Iterative Cross-Assembly of Sponge-Associated Viromes

Clustering shows the distance between viral metagenomes based on the fraction of cross-assembled contigs between all sample pairs. Comparing topology against 1,000 random trees indicates significant separation of environments (sponge versus seawater) and sponge species (p value $\leq$ 0.001) but not sponge tissues (p value = 0.991). Heatmap shows the relative abundances of viral contigs and their grouping into prevalence groups as detailed in the STAR Methods section. Color scheme is based on *Z* score distribution across samples from low (blue) to high (red).

populations that could be annotated at the family level, representing mainly bacteriophages of the Caudovirales families, Siphoviridae, Myoviridae, and Podoviridae (Figure S2B). Rarefaction analysis indicated that more sponge viral diversity remains to be discovered, as the curve has not reached saturation (Figure 1C). These observations, combined with the limited taxonomic overlap with other marine environments (Figure 1C), led us to the conclusion that sponges represent distinct niches for viruses with potential for previously undescribed functions.

Unique Viromes in Neighboring Sponges

Viral communities of neighboring sponges were individually unique, host species specific, and different from environmental seawater (Figure 2). This is indicated by the fact that viral community profiles grouped per sponge species (p value < 0.001, consistency value = 0.907) and were distinct from adjacent viroplankton (p value < 0.001, consistency value = 0.973). Variation in viral community composition within a given sponge species was mainly on the level of sponge individuals (p value < 0.001, consistency value = 0.679), rather than tissue-specific signatures (p value = 0.991, consistency value = 0.534). These observations based on the fraction of cross contigs between the sample pairs (detailed in STAR Methods) showed high concordance with results from hierarchical clustering of abundance profiles (Figure S4). We further explored viral populations by conceptualizing viral prevalence groups (see STAR Methods for details). These were the generalists (prevalent in more than one sponge species or seawater), specialists (prevalent in one sponge species or seawater), individualists (detected in only one individual but both tissues), and intermediates (not falling into the above definitions) (Figure 2). Notably, even though we obtained the samples from neighboring sponges of each sponge species at the same time point, individualists, at 34.8% (1,560 of 4,484 BCvir contigs), represented the largest virome group in our study. Furthermore, individualists were the second most abundant prevalence group in the dataset, indicating that these were not rare members of the community. In contrast, a minor fraction of BCvir population contigs were generalists being prevalent in all (n = 10) or multiple (n = 262) sample types (species and/or seawater). Specialists, with prevalence in one of the sponge species or seawater, were 27.9% (1,249 of 4,484). The intermediates, although not further categorized, still contain a species-specific pattern. Because efforts were made to minimize environmental variations by sampling in close spatial and temporal proximity, we conclude that the sponge individuals each have a unique viral fingerprint.

Symbiont Phage Protein Aids Bacteria in Eukaryote Immune Evasion

To identify factors that might improve fitness of the phages, we queried BCvir populations for auxiliary genes (Roux et al.,

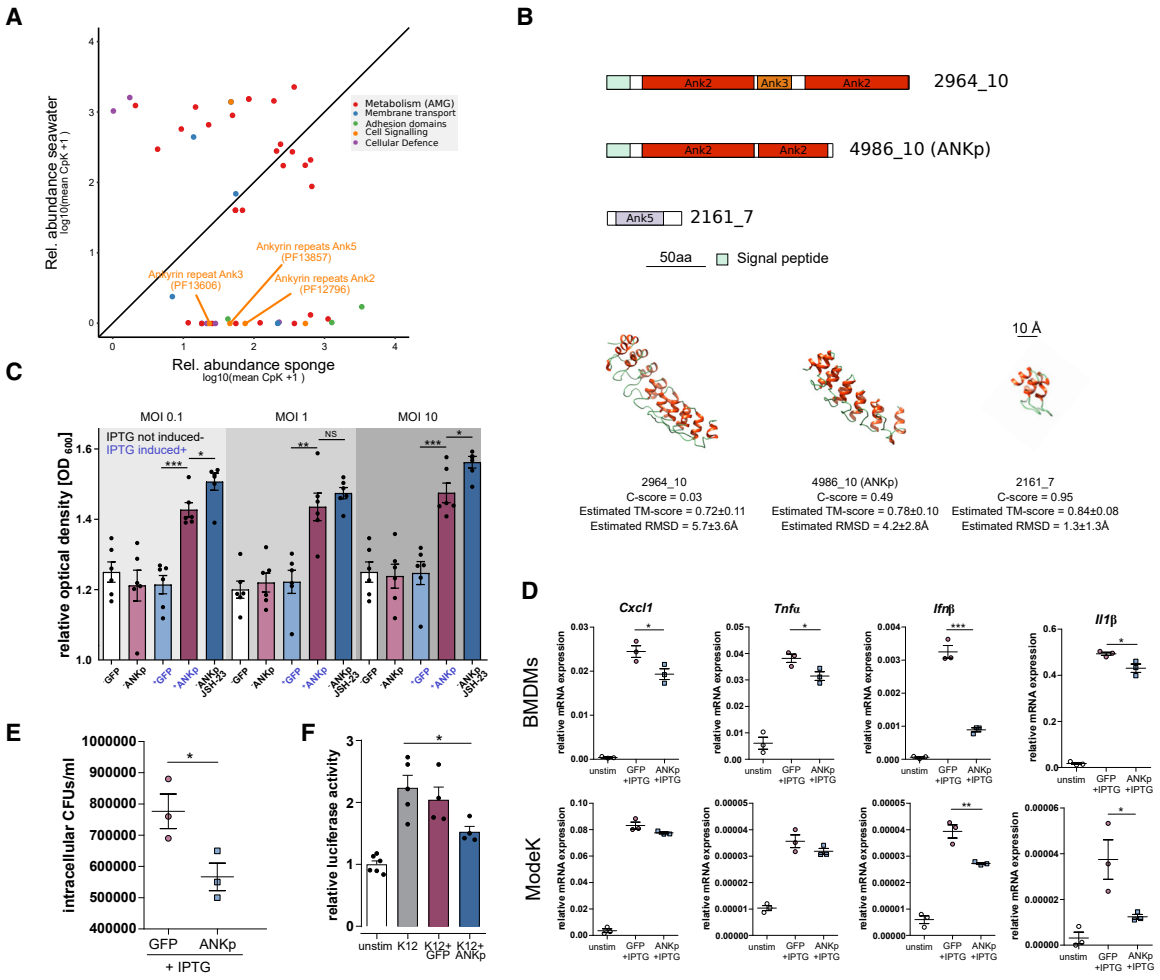


Figure 3. Symbiont Phage ANKp Reduces Phagocytosis and Immune Response of Eukaryote Cells toward Bacteria

(A) Relative abundance of auxiliary viral functions in sponge versus seawater (raw data and further host association factors available in Data S1).

(B) Domain architecture of ankyrin repeat encoding genes from sponge-enriched phages and representative protein models, approximated using I-TASSER. Template modeling (TM) score measures protein similarity and the confidence (C)-score prediction accuracy (Roy et al., 2010). RMSD indicates model root-mean-square deviation.

(C) Growth kinetics of *E. coli* K12 that are challenged with murine bone-marrow-derived macrophages (BMDMs) upon expression of recombinant ANKp. Plus^{+} and minus^{-} indicate treatments with and without IPTG induction of protein expression. Strains expressing GPF were used as the negative control and the NF-κB inhibitor JSH-23 was used as the positive control for immune inhibition. Data are presented as the mean ± SEM of three independent experiments (each with n = 6).

(D) Expression levels of pro-inflammatory cytokines in macrophages (BMDMs) and intestinal epithelial cells ModeK upon infection with ANKp expressing *E. coli*.

(E) Gentamycin protection assay reveals that ANKp expression leads to a reduced number of intracellular bacteria.

(F) NF-κB activation in ModeK with dual-luciferase assay.

Data are presented as the mean ± SEM of at least three independent experiments. Statistical significance between treatments was determined by two-tailed unpaired Student's t tests with *p < 0.05, **p < 0.01, and ***p < 0.001.

2016; Data S1). We then extended our search for cellular membrane transporters, adhesins, defense systems, and cellular signal molecules owing to their potential relevance in a symbiosis context. We were surprised to find Ankyrin repeat domains (ANKs), discussed modulators of eukaryote-prokaryote interaction (Nguyen et al., 2014), to be encoded on sponge-associated phages (Figure 3A). These ANK-encoding phages (BCvir 2964, BCvir 2161, and BCvir 4986), which we will call Ankyphage 1, 2, and 3 hereinafter (Data S1; Ankyphage annotation), recruited reads from 12 of 32 sponge viromes from both pinacoderm and mesohyl tissues but were absent in seawater. All three Ankyphages fall in the category "intermediates" (Figure 2). Furthermore, Ankyphages were in the top 75th percentile of most abundant viruses detected in *Aplysina* and *Chondrosia*. To ensure that Ankyphage sequences are indeed phage, we confirmed their phylogenetic placement among bacteriophages based on capsid alignments (Figures S3A and S3B) and ensured on the same contigs the presence of further phage domains (Figure S3C), such as phage terminase (PF03354), phage portal protein (PF04860), and phage P22 coat protein (PF11651). The presence of *ANK* in CsCl-purified virus particles was confirmed by Sanger sequencing. Notably, the domain architecture of two Ankyphage ANKs comprised N-terminal signal peptides but no transmembrane domains (Figure 3B). This suggests that

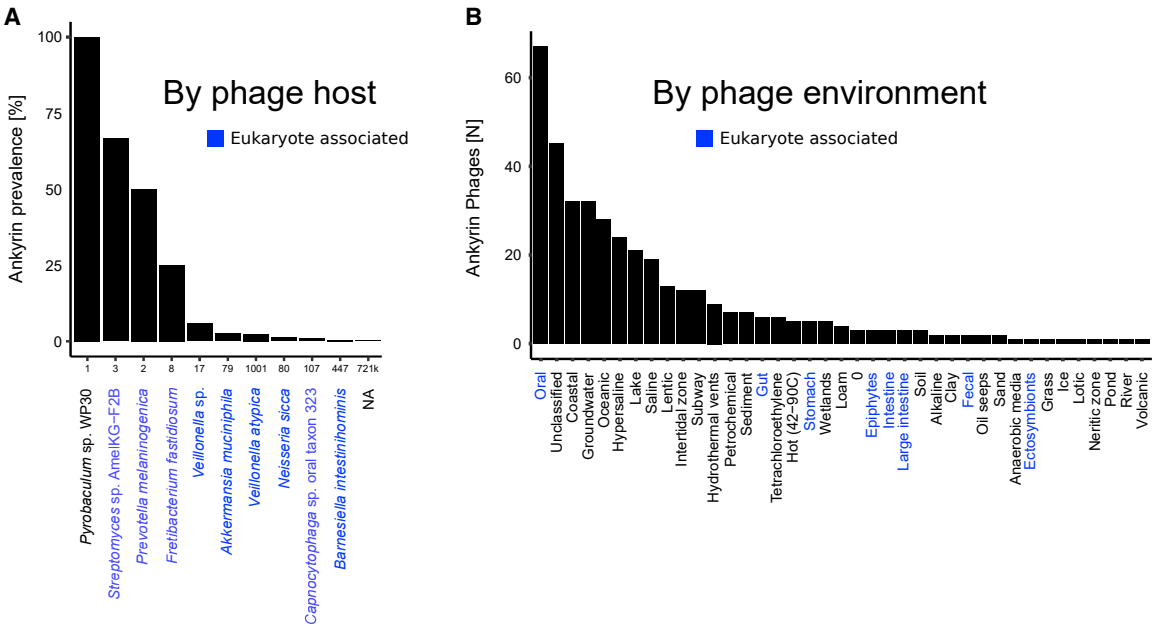


Figure 4. Global Distribution of Phage ANKs as Deposited in the IMGvr Database

(A) Prevalence of ankyrins in phages per predicted prokaryotic host (as deposited in IMGvr database metadata). Numbers on the x axis indicate the total size of phage genomes or contigs per host taxon.

(B) Number of phages with ANK domains per environment. Prokaryotes or environments with eukaryote host associations are highlighted in blue. The data are based on IMGvr database (Paez-Espino et al., 2019) screening for ankyrin Pfam signatures Ank (PF00023), Ank_2 (PF12796), Ank_3 (PF13606), Ank_4 (PF13637), and Ank_5 (PF13857) using InterPro (see STAR Methods).

these ANKs are secreted from a phage-infected bacterial cell where Ankyphages may be integrated as prophages or pseudolysogens with stable episome.

We then synthesized Ankyphage 3 ANK protein (ANKp) and assayed its impact on the interaction between bacteria and macrophages. Murine cell lines were chosen for the lack of an experimentally tractable model for sponge-microbe interactions (Pita et al., 2016). Murine macrophages functionally resemble certain sponge cells (archaeocytes) that are single, amoeboid, and phagocytotically active cells, which patrol throughout the sponge matrix. Notably, Ankyphages were found as well throughout all sponge tissues. Moreover, the major elements involved in mammalian immune signaling were found to be present in sponges (Pita et al., 2018; Riesgo et al., 2014). When *E. coli* was pre-incubated with purified ANKp, the bacterium survived in significantly higher abundances upon predatory pressure from murine macrophages than GFP controls (Figures S3F and S3D). This effect was dose dependent (1 μM > 100 nM ANKp) and reproducible for different bacteria to macrophage ratios (multiplicity of infection (MOI) 0.1, 1, and 10). Gentamycin protection assays, which allowed us to quantify the intracellular bacteria fraction, showed that increased survival of ANKp pre-incubated *E. coli* was paralleled with a decreased number of intracellular bacteria expressing the protein (p value = 0.0414, t = 2.963, and df = 4) (Figure 3E). This indicates that ANKp-mediated bacterial survival is facilitated by decreased macrophage phagocytosis rates. In support of these findings, when ANKp was directly expressed by *E.coli*, this resulted in a significantly increased survival rate of *E.coli* upon exposure to macrophages, showing that ANKp is functional when secreted (Figures 3C and S3E). To ensure that ANKp protein had no toxic effect on one of the players, we performed bacterial growth experiments in culture and on plates and for the eukaryotic cell line MTS assays upon protein exposure showing low cytotoxicity (Figures S3G–S3I). The same growth experiment exercised with pre-incubated *Bacillus subtilis*, a Gram-positive representative, was consistent with ANKp-mediated bacterial survival during macrophage challenge (Figure S3C).

On the side of the macrophage, ANKp synthesis by *E. coli* led to a reduced expression of pro-inflammatory cytokines upon bacterial exposure (Figure 3D). Specifically, this included a reduction in tumor necrosis factor alpha (*TNF-α*), *Cxcl1*, and *Ifn1*. To independently validate phage ANKp-mediated eukaryote immune suppression and to extend the analysis to a further eukaryotic cell type, we performed an NF-κB-dependent firefly luciferase assay on murine gut endothelial ModeK cells. In line with previous results, the NF-κB response, a central hub of eukaryote immunity (Li and Verma, 2002), was downregulated when ModeK cells were exposed to ANKp expressing *E. coli* (Figure 3F). In summary, this shows that ANKp modulates the eukaryote response to bacteria by downregulating pro-inflammatory signaling along with reduced phagocytosis rates.

To investigate whether phage ankyrins are more common in nature, we extended homology searches to various other viral databases including IMGvr (Paez-Espino et al., 2019) (July 2018 release). We identified an abundant ANK-encoding virus in the Great Barrier Reef sponge *Amphimedon queenslandica* (Laffy et al., 2018) (Data S1; Ankyphage annotation). Furthermore, we identified ANKs in 418 predicted phage contigs deposited in IMGvr (Paez-Espino et al., 2019) (Figure 4). Notably,

Ankyphages were obtained from host-associated environments such as the human oral cavity (n = 67 phages), gut (n = 6), stomach (n = 5), or rhizosphere (n = 4) but also from aquatic environments ranging from marine to groundwater.

DISCUSSION

We discovered an astonishingly intimate association between viruses and marine sponges. Altogether 491 yet unexplored VCs were identified in sponges that delineate genus-level taxonomy (Lima-Mendez et al., 2008; Roux et al., 2015). This compares to 1,843 VCs from the Tara Oceans and other large-scale expeditions focusing on viroplankton. Although the applied genome-based network clustering approach cannot replace well-curated taxonomy (King et al., 2018), the results of us and others (Paez-Espino et al., 2016) indicate that marine animals indeed represent distinct niches of viral diversity. Consequently, future efforts to capture more host-associated environments have high potential to add up to the increasingly understood planktonic virosphere of surface waters (Coutinho et al., 2017; Gregory et al., 2019; Roux et al., 2016). Because animal microbiomes are highly species specific (Hacquard et al., 2015; Moeller et al., 2016; Thomas et al., 2016) and the virome depends on the microbiome, we further expect species-specific viral communities to rule in nature. This is supported by our sympatric sponge species, each holding characteristic viral communities (Figure 2). Systematic studies in other sponges (Laffy et al., 2018) but also in *Hydra* (Grasis et al., 2014) and insects, where signatures of phylosymbiosis were observed (Leigh et al., 2018a), are supportive for species-specific viromes in animals.

An even higher level of association of viruses with their animal hosts was identified by our nested sampling approach (Figure 2). This revealed that a considerable part of the virome signatures was driven by viruses that were unique to sponge individuals but not to tissues (Figure 2). Inter-individual differences were also the largest source of variance in the viromes of humans (Abeles et al., 2014; Moreno-Gallego et al., 2019). Our report systematically extends individual virome signatures to marine sponges and adds to evidence of virome individuality in other marine animals (Leigh et al., 2018b; Orosco and Lluisma, 2017). The high degree of individuality in sponge viromes was surprising considering the constant filtration activity of sponges (Taylor et al., 2013). The diversifying forces accounting for individual viromes may be asynchronous temporal fluctuations between sponge individuals following delayed Lotka-Volterra-like dynamics (Parsons et al., 2012) or an independent diversification from a source pool (Enav et al., 2018). To resolve this, further studies capturing the temporal dynamics of environmental host-associated systems will provide valuable information.

The discovery of auxiliary ankyrin repeats (ANKs) in a previously undescribed group of sponge-associated phages, that we term Ankyphages, raised our special interest due to (1) their protein architecture, which indicates its secretion from the virocell; (2) their role as hubs of diverse protein-protein interactions, including functions in cellular signaling; and (3) their broader prevalence and abundance in the symbiotic context while lacking in nearby seawater. We reasoned that phage-encoded ANKs might increase the fitness of the carrying phages in the context of the sponge holobiont. ANK repeats are widespread in all domains of cellular life (Jernigan and Bordenstein, 2014), but reports of ANKs in the world of phages are rare. A notable exception is PRANC domains (Pox protein repeats of ankyrin CTD), which are ANK homologs of poxviruses discovered in prophages of Wolbachia (Bordenstein and Bordenstein, 2016; Wu et al., 2004). Its placement in a conserved eukaryote association module indicates its functioning in a eukaryotic context (Bordenstein and Bordenstein, 2016). In the bacterial world, ANKs were demonstrated to modulate interaction between species and even across kingdoms (Lambert et al., 2015; Wong et al., 2017). A notable example is a previous study on sponges, where chromosomally encoded ANKs from an uncultivated gammaproteobacterium seemed to modulate amoebal phagocytosis (Nguyen et al., 2014), even though the underlying mechanisms on the eukaryote side are largely unclear.

Our *in vitro* experiments show that phage ANKp undermined eukaryote immune response toward bacteria and facilitated bacteria-eukaryote coexistence by reduced phagocytosis rates. To the best of our knowledge, a secreted phage protein shown to downregulate eukaryote immune response has not been previously described. This finding has important implications from a symbiotic perspective. The reduction of predatory pressure from the eukaryote host represents a selective advantage for the symbiotic lifestyle of Ankyphage-infected bacteria as compared to strains missing this trait (see Graphical Abstract). Eukaryote immune evasion by phage-mediated lysogenic conversion is an emerging field of research that is currently best studied in opportunistic pathogens (Van Belleghem et al., 2018). Mechanisms range from phage-mediated reshaping of methicillin-resistant *Staphylococcus aureus* (MRSA) cell wall glycosylation to evade host immunity (Gerlach et al., 2018), to phage RNA of the *Pseudomonas aeruginosa* phage Pf4 that downregulates eukaryote inflammatory response and at the same time is taming for non-invasive infection (Secor et al., 2017; Sweere et al., 2019). We are aware that the choice for the experimentally more approachable murine model can only be a proxy for processes yet to be observed in sponges. However, consistent signals in *E. coli* and *B. subtilis* and the tested eukaryote cell type (macrophages and an epithelial cell line), might indicate a more widely distributed conserved mode of action. This is fueled by our public database screenings where we found phage-encoded ANKs in other eukaryote-associated environments, such as phages inhabiting human cavities (oral, stomach, or gut) representing a promising field for future research. In summary, our study highlights the novel diversity, intimate association, and tripartite interplay between phages, symbionts, and the eukaryote host. Importantly, we identify and characterize a phage-derived protein that can manipulate the immune interaction between eukaryotes and microbiota.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chom.2019.08.019.

ACKNOWLEDGMENTS

We acknowledge funding by the DFG CRC1182 to U.H. (TPB1), T.L. (TPA4), P.R. (TPC2), and A.K. (TPC3). M.T.J. was supported by a grant of the German Excellence Initiative to the Graduate School of Life Sciences, University of Würzburg, and the Young Investigator Award of CRC1182. S.M.M. was supported by the Studienstiftung des Deutschen Volkes. We thank Laura Rix (GEOMAR), Rafel Coma (CEAB-CSIC), and Berta Pintó for sponge sampling.

AUTHOR CONTRIBUTIONS

M.T.J. designed and performed the experiments, analyzed the data, prepared the figures and tables, and wrote the paper; L.P. and M.R. were involved in the planning and execution of the sponge field sampling; T.L. purified the viral particles and extracted the nucleotides; S.T.S. and P.R. performed cell culture experiments and immunoassays; M.T.J. and S.M.M. performed the microscopy under the supervision of C.S.; B.E.D., A.K., and K.A. advised the bioinformatic analyses and commented on the analysis strategy; U.H. helped in the experimental design and data interpretation and reviewed drafts of the paper. All authors edited and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 8, 2019
Revised: July 22, 2019
Accepted: August 30, 2019
Published: September 24, 2019

REFERENCES

Abeles, S.R., Robles-Sikisaka, R., Ly, M., Lum, A.G., Salzman, J., Boehm, T.K., and Pride, D.T. (2014). Human oral viruses are personal, persistent and gender-consistent. ISME J. *8*, 1753–1767.

Barr, J.J., Auro, R., Furlan, M., Whiteson, K.L., Erb, M.L., Pogliano, J., Stotland, A., Wolkowicz, R., Cutting, A.S., Doran, K.S., et al. (2013). Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc. Natl. Acad. Sci. USA *110*, 10771–10776.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Betts, A., Gray, C., Zelek, M., MacLean, R.C., and King, K.C. (2018). High parasite diversity accelerates host adaptation and diversification. Science *360*, 907–911.

Bordenstein, S.R., and Bordenstein, S.R. (2016). Eukaryotic association module in phage WO genomes from Wolbachia. Nat. Commun. *7*, 13155.

Breitbart, M., Bonnain, C., Malki, K., and Sawaya, N.A. (2018). Phage puppet masters of the marine microbial realm. Nat. Microbiol. *3*, 754–766.

Conceição-Neto, N., Zeller, M., Lefrère, H., De Bruyn, P., Beller, L., Deboutte, W., Yinda, C.K., Lavigne, R., Maes, P., Van Ranst, M., et al. (2015). Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Sci. Rep. *5*, 16532.

Coutinho, F.H., Silveira, C.B., Gregoracci, G.B., Thompson, C.C., Edwards, R.A., Brussaard, C.P.D., Dutilh, B.E., and Thompson, F.L. (2017). Marine viruses discovered via metagenomics shed light on viral strategies throughout the oceans. Nat. Commun. *8*, 15955.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. J. Veg. Sci. *14*, 927–930.

Dutilh, B.E., Schmieder, R., Nulton, J., Felts, B., Salamon, P., Edwards, R.A., and Mokili, J.L. (2012). Reference-independent comparative metagenomics using cross-assembly: crAss. Bioinformatics *28*, 3225–3231.

Edwards, R.A., Vega, A.A., Norman, H.M., Ohaeri, M., Levi, K., Dinsdale, E.A., Cinek, O., Aziz, R.K., McNair, K., Barr, J.J., et al. (2019). Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. Nat. Microbiol.

Enav, H., Kirzner, S., Lindell, D., Mandel-Gutfreund, Y., and Béjà, O. (2018). Adaptation to sub-optimal hosts is a driver of viral diversification in the ocean. Nat. Commun. *9*, 4698.

Enright, A.J., Van Dongen, S., and Ouzounis, C.A. (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. *30*, 1575–1584.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ *3*, e1319.

Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., Melamed, S., Leavitt, A., Savidor, A., Albeck, S., et al. (2017). Communication between viruses guides lysis–lysogeny decisions. Nature *541*, 488–493.

Gerlach, D., Guo, Y., De Castro, C., Kim, S.H., Schlatterer, K., Xu, F.F., Pereira, C., Seeberger, P.H., Ali, S., Codée, J., et al. (2018). Methicillin-resistant Staphylococcus aureus alters cell wall glycosylation to evade immunity. Nature *563*, 705–709.

Grasis, J.A., Lachnit, T., Anton-Erxleben, F., Lim, Y.W., Schmieder, R., Fraune, S., Franzenburg, S., Insua, S., Machado, G., Haynes, M., et al. (2014). Species-specific viromes in the ancestral holobiont Hydra. PLoS One *9*, e109952.

Grazziotin, A.L., Koonin, E.V., and Kristensen, D.M. (2017). Prokaryotic Virus Orthologous Groups (pVOGs): a resource for comparative genomics and protein family annotation. Nucleic. Acids. Res. *45*, D491–D498.

Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Cruaud, C., et al. (2019). Marine DNA viral macro- and microdiversity from Pole to Pole. Cell *177*, 1109–1123.e14.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. Bioinformatics *29*, 1072–1075.

Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A.C., Dangl, J.L., Knight, R., Ley, R., et al. (2015). Microbiota and host nutrition across plant and animal kingdoms. Cell Host Microbe *17*, 603–616.

Horn, H., Slaby, B.M., Jahn, M.T., Bayer, K., Moitinho-Silva, L., Förster, F., Abdelmohsen, U.R., and Hentschel, U. (2016). An enrichment of CRISPR

and other defense-related features in marine sponge-associated microbial metagenomes. Front. Microbiol. *7*, 1751.

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.

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.

Jernigan, K.K., and Bordenstein, S.R. (2014). Ankyrin domains across the tree of life. PeerJ *2*, e264.

John, S.G., Mendez, C.B., Deng, L., Poulos, B., Kauffman, A.K.M., Kern, S., Brum, J., Polz, M.F., Boyle, E.A., and Sullivan, M.B. (2011). A simple and efficient method for concentration of ocean viruses by chemical flocculation. Environ. Microbiol. Rep. *3*, 195–202.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: Genome-scale protein function classification. Bioinformatics *30*, 1236–1240.

Kaser, A., Lee, A.H., Franke, A., Glickman, J.N., Zeissig, S., Tilg, H., Nieuwenhuis, E.E., Higgins, D.E., Schreiber, S., Glimcher, L.H., et al. (2008). XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. Cell *134* (5), 743–756.

Keen, E.C., and Dantas, G. (2018). Close encounters of three kinds: bacteriophages, commensal bacteria, and host immunity. Trends Microbiol. *26*, 943–954.

King, A.M.Q., Lefkowitz, E.J., Mushegian, A.R., Adams, M.J., Dutilh, B.E., Gorbalenya, A.E., Harrach, B., Harrison, R.L., Junglen, S., Knowles, N.J., et al. (2018). Changes to taxonomy and the International code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2018). Arch. Virol. *163*, 2601–2631.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. *305* (3), 567–580.

Kronheim, S., Daniel-Ivad, M., Duan, Z., Hwang, S., Wong, A.I., Mantel, I., Nodwell, J.R., and Maxwell, K.L. (2018). A chemical defence against phage infection. Nature *564*, 283–286.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. *5*, R12.

Lachnit, T., Thomas, T., and Steinberg, P. (2015). Expanding our understanding of the seaweed holobiont: RNA viruses of the red alga Delisea pulchra. Front. Microbiol. *6*, 1489.

Laffy, P.W., Wood-Charlson, E.M., Turaev, D., Jutz, S., Pascelli, C., Botté, E.S., Bell, S.C., Peirce, T.E., Weynberg, K.D., van Oppen, M.J.H., et al. (2018). Reef invertebrate viromics: diversity, host specificity and functional capacity. Environ. Microbiol. *20*, 2125–2141.

Lambert, C., Cadby, I.T., Till, R., Bui, N.K., Lerner, T.R., Hughes, W.S., Lee, D.J., Alderwick, L.J., Vollmer, W., Sockett, R.E., et al. (2015). Ankyrin-mediated self-protection during cell invasion by the bacterial predator Bdellovibrio bacteriovorus. Nat. Commun. *6*, 8884.

Leigh, B.A., Bordenstein, S.R., Brooks, A.W., Mikaelyan, A., and Bordenstein, S.R. (2018a). Finer-scale phylosymbiosis: insights from insect viromes. mSystems *3*, e00131-18.

Leigh, B.A., Djurhuus, A., Breitbart, M., and Dishaw, L.J. (2018b). The gut virome of the protochordate model organism, Ciona intestinalis subtype A. Virus Res. *244*, 137–146.

Letunic, I., and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. *47*, W256–W259.

Li, Q., and Verma, I.M. (2002). NF-kappaB regulation in the immune system. Nat. Rev. Immunol. *2*, 725–734.

Lima-Mendez, G., Van Helden, J., Toussaint, A., and Leplae, R. (2008). Reticulate representation of evolutionary and functional relationships between phage genomes. Mol. Biol. Evol. *25*, 762–777.

López-Pérez, M., Haro-Moreno, J.M., Gonzalez-Serrano, R., Parras-Moltó, M., and Rodriguez-Valera, F. (2017). Genome diversity of marine phages recovered from Mediterranean metagenomes: size matters. PLoS Genet. *13*, e1007018.

Marston, M.F., Pierciey, F.J., Shepard, A., Gearin, G., Qi, J., Yandava, C., Schuster, S.C., Henn, M.R., and Martiny, J.B.H. (2012). Rapid diversification of coevolving marine Synechococcus and a virus. Proc. Natl. Acad. Sci. USA *109*, 4544–4549.

Moeller, A.H., Caro-Quintero, A., Mjungu, D., Georgiev, A.V., Lonsdorf, E.V., Muller, M.N., Pusey, A.E., Peeters, M., Hahn, B.H., and Ochman, H. (2016). Cospeciation of gut microbiota with hominids. Science *353*, 380–382.

Moreno-Gallego, J.L., Chou, S.P., Di Rienzi, S.C., Goodrich, J.K., Spector, T.D., Bell, J.T., Youngblut, N.D., Hewson, I., Reyes, A., and Ley, R.E. (2019). Virome diversity correlates with intestinal microbiome diversity in adult monozygotic twins. Cell Host Microbe *25*, 261–272.e5.

Nguyen, M.T., Liu, M., and Thomas, T. (2014). Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. Mol. Ecol. *23*, 1635–1645.

Nielsen, H. (2017). Predicting Secretory Proteins with SignalP. Methods Mol. Biol. *1611*, 59–73.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017). metaSPAdes: a new versatile metagenomic assembler. Genome Res. *27*, https://doi.org/10.1101/gr.213959.116.

Orosco, F.L., and Lluisma, A.O. (2017). Variation in virome diversity in wild populations of Penaeus monodon (Fabricius 1798) with emphasis on pathogenic viruses. Virusdisease *28*, 262–271.

Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N.N., and Kyrpides, N.C. (2016). Uncovering earth's virome. Nature *536*, 425–430.

Paez-Espino, D., Roux, S., Chen, I.A., Palaniappan, K., Ratner, A., Chu, K., Huntemann, M., Reddy, T.B.K., Pons, J.C., Llabrés, M., et al. (2019). IMG/VR v.2.0: an integrated data management and analysis system for cultivated and environmental viral genomes. Nucleic Acids Res. *47*, D678–D686.

Parsons, R.J., Breitbart, M., Lomas, M.W., and Carlson, C.A. (2012). Ocean time-series reveals recurring seasonal patterns of virioplankton dynamics in the northwestern Sargasso Sea. ISME J. *6*, 273–284.

Pita, L., Fraune, S., and Hentschel, U. (2016). Emerging sponge models of animal-microbe symbioses. Front. Microbiol. *7*, 2102.

Pita, L., Hoeppner, M.P., Ribes, M., and Hentschel, U. (2018). Differential expression of immune receptors in two marine sponges upon exposure to microbial-associated molecular patterns. Sci. Rep. *8*, 16081.

Podell, S., Blanton, J.M., Neu, A., Agarwal, V., Biggs, J.S., Moore, B.S., and Allen, E.E. (2019). Pangenomic comparison of globally distributed Poribacteria associated with sponge hosts and marine particles. ISME J. *13*, 468–481.

Riesgo, A., Farrar, N., Windsor, P.J., Giribet, G., and Leys, S.P. (2014). The analysis of eight transcriptomes from all Poriferan classes reveals surprising genetic complexity in sponges. Mol. Biol. Evol. *31*, 1102–1120.

Rohwer, F. (2003). Global phage diversity. Cell *113*, 141.

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E.V., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Lavigne, R., Brister, J.R., Varsani, A., et al. (2019). Minimum information about an uncultivated virus genome (MIUViG). Nat. Biotechnol. *37*, 29–37.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., Poulos, B.T., Solonenko, N., Lara, E., Poulain, J., et al. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. Nature *537*, 689–693.

Roux, S., Emerson, J.B., Eloe-Fadrosh, E.A., and Sullivan, M.B. (2017). Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ *5*, e3817.

Roux, S., Hallam, S.J., Woyke, T., and Sullivan, M.B. (2015). Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. Elife *4*.

Roux, S., Krupovic, M., Debroas, D., Forterre, P., and Enault, F. (2013). Assessment of viral community functional potential from viral metagenomes

may be hampered by contamination with cellular sequences. Open Biol. *3*, 130160.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. *5*, 725–738.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., et al. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. *75* (23), 7537–7541.

Secor, P.R., Michaels, L.A., Smigiel, K.S., Rohani, M.G., Jennings, L.K., Hisert, K.B., Arrigoni, A., Braun, K.R., Birkland, T.P., Lai, Y., et al. (2017). Filamentous bacteriophage produced by Pseudomonas aeruginosa alters the inflammatory response and promotes noninvasive infection in vivo. Infect. Immun. *85*, e00648-16.

Shen, A., Lupardus, P.J., Morell, M., Ponder, E.L., Sadaghiani, A.M., Garcia, K.C., and Bogyo, M. (2009). Simplified, enhanced protein purification using an inducible, autoprocessing enzyme tag. PLoS One *4*, e8119.

Slaby, B.M., Hackl, T., Horn, H., Bayer, K., and Hentschel, U. (2017). Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization. ISME J. *11*, 2465–2478.

Suttle, C.A. (2007). Marine viruses–major players in the global ecosystem. Nat. Rev. Microbiol. *5*, 801–812.

Sweere, J.M., Van Belleghem, J.D., Ishak, H., Bach, M.S., Popescu, M., Sunkari, V., Kaber, G., Manasherob, R., Suh, G.A., Cao, X., et al. (2019). Bacteriophage trigger antiviral immunity and prevent clearance of bacterial infection. Science *363*.

Taylor, M.W., Tsai, P., Simister, R.L., Deines, P., Botte, E., Ericson, G., Schmitt, S., and Webster, N.S. (2013). 'Sponge-specific' bacteria are widespread (but rare) in diverse marine environments. ISME J. *7*, 438–443.

Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., Olson, J.B., Erwin, P.M., López-Legentil, S., Luter, H., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. Nat. Commun. *7*, 11870.

Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., and Rohwer, F. (2009). Laboratory procedures to generate viral metagenomes. Nat. Protoc. *4*, 470–483.

Van Belleghem, J.D., Dąbrowska, K., Vaneechoutte, M., Barr, J.J., and Bollyky, P.L. (2018). Interactions between bacteriophage, bacteria, and the mammalian immune system. Viruses *11*, 10.

Weisz, J.B., Lindquist, N., and Martens, C.S. (2008). Do associated microbial abundances impact marine demosponge pumping rates and tissue densities? Oecologia *155*, 367–376.

Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. *64*, 69–114.

Wong, K., Perpich, J.D., Kozlov, G., Cygler, M., Abu Kwaik, Y., and Gehring, K. (2017). Structural mimicry by a bacterial F Box effector hijacks the host ubiquitin-proteasome system. Structure *25*, 376–383.

Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., et al. (2004). Phylogenomics of the reproductive parasite Wolbachia pipientis wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol. *2*, E69.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
D3I1O XP Rabbit mAb	Cell Signaling	Cat#12698; RRID: AB_2744546
Bacterial and Virus Strains		
Bacillus subtilis	DSMZ	DSM 10
Escherichia coli K12	DSMZ	DSM 498
BL21(DE3) *E. coli*	New England Biolabs	Cat#C2527I
Biological Samples		
Sponge: *Agelas oroides*	This manuscript	N/A
Sponge: *Aplysina aerophoba*	This manuscript	N/A
Sponge: *Chondrosia reniformis*	This manuscript	N/A
Sponge: *Petrosia ficiformis*	This manuscript	N/A
Chemicals, Peptides, and Recombinant Proteins		
Polyvinylpolypyrrolidone (PVPP)	Sigma-Aldrich	Cat#77627 CAS Number: 9003-39-8
Cesium chloride	Fisher Scientific	Cat#10648783 CAS Number: 7647-17-8
IPTG	Sigma-Aldrich	Cat#I6758 CAS Number: 367-93-1
JSH-23	Sigma-Aldrich	Cat#J4455CAS Number: 749886-87-1
Critical Commercial Assays		
FuGENE 6 Transfection Reagent	Promega	Cat#E2691
Maxima H Minus First Strand cDNA Synthesis kit	Thermo Scientific	Cat#K1682
Nextera XT DNA Library Preparation Kit	Illumina	Cat#FC-131-1096
pNF-κB-Luc	Clontech	N/A
Ni-NTA Fast Start Kit	Qiagen	Cat#30600
pRL-TK	Clontech	N/A
RNeasy Mini Kit	Qiagen	Cat#74106
TaqMan Gene Expression Master Mix	Applied Biosystems	Cat#4369016
Whole Transcriptome Amplification Kit 2	Sigma-Aldrich	Cat#WTA2-50RXN
Deposited Data		
Viromics raw data	This manuscript	BioProject: PRJNA522695
Microbiota 16S rDNA gene sequences	This manuscript	BioProject: PRJNA522695
Experimental Models: Cell Lines		
MODE-K cells	Kaser et al., 2008	N/A
Experimental Models: Organisms/Strains		
Mice: female C57BL/6	own breeding	N/A
Oligonucleotides		
Cxcl1	TaqMan Gene Expression Assays for mouse,Life Technologies	Mm00433859_m1
Gapdh	TaqMan Gene Expression Assays for mouse,Life Technologies	Mm99999915_g1
Ifnb1	TaqMan Gene Expression Assays for mouse,Life Technologies	Mm00439552_s1
Tnfα	TaqMan Gene Expression Assays for mouse,Life Technologies	Mm00443258_m1
Recombinant DNA		
pET22b-GFP-CPDSal	Shen et al., 2009	Addgene Cat# 38257 RRID:Addgene_38257
pET22b-ANKp	This manuscript	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Anvi'o v.2.1.1	Eren et al., 2015	http://merenlab.org/software/anvio/
BBMap v.37.7	Bushnell B.	https://sourceforge.net/projects/bbmap/
CAT	Bas E. Dutilh	https://github.com/dutilh/CAT
cluster_screener	This manuscript	https://github.com/MartinTJahn/cluster_screener
crAss-Tool	Dutilh et al., 2012	http://crass.sourceforge.net
Prism 6 software v.6	GraphPad	https://www.graphpad.com/scientific-software/prism/
Markov cluster algorithm (MCL)	Enright et al., 2002	https://micans.org/mcl/
QUAST v5.0.2	Gurevich et al., 2013	http://quast.sourceforge.net/quast.html
InterProScan v5.27-66.0	Jones et al., 2014	https://www.ebi.ac.uk/interpro/download.html
I-TASSER v.5.1	Roy et al., 2010	https://zhanglab.ccmb.med.umich.edu/I-TASSER/
iTOL v.4	Letunic and Bork, 2019	https://itol.embl.de/
LMCLUST	This manuscript	https://github.com/kseniaarkhipova/LMCLUST
metaSPAdes v.3.11.1	Nurk et al., 2017	https://github.com/ablab/spades/releases
mothur v.1.39.5	Schloss et al., 2009	https://www.mothur.org/wiki/Main_Page
PRODIGAL v2.6.3	Hyatt et al., 2010	https://github.com/hyattpd/Prodigal
RedRed	Ksenia Arkhipova	https://github.com/kseniaarkhipova/RedRed
SignalP v4.1f	Nielsen, 2017	http://www.cbs.dtu.dk/services/SignalP/
Tecan i-control v.1.9	Tecan	https://lifesciences.tecan.com/
TMHMM 2.0c	Krogh et al., 2001	http://www.cbs.dtu.dk/services/TMHMM/
treestats	Rob Edwards	https://github.com/linsalrob/crAssphage/tree/master/bin
VirSorter v.1.0.3	Roux et al., 2015	https://github.com/simroux/VirSorter
7900HT Fast Real-Time PCR Software v.2.4.1	Applied Biosystems	https://www.thermofisher.com

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ute Hentschel (uhentschel@geomar.de). All unique reagents generated in this study are available from the Lead Contact without restriction.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Generation of Bone Marrow-Derived Macrophages (BMDMs)

C57BL/6 mice (10-14 weeks old) were killed, femur and tibia were dissected, and bone marrow was isolated under sterile conditions. Haematopoietic stem cells differentiated into BMDMs by incubating for 7 days in BMDM medium (1:1 SFM:DMEM, Gibco) supplemented with 10 % FCS (Biochrom), 1 % penicillin/streptomycin (Gibco), 1 % amphotericin B (Gibco), and 20 ng/mL macrophage colony-stimulating factor (mCSF, Immunotools). All animal experiments were approved by the Animal Investigation Committee of the University Hospital Schleswig-Holstein (Campus Kiel, Germany; acceptance no.: V242-7224.121-33) and were performed according the relevant guidelines and regulations.

ModeK Cells

ModeK cells were kindly provided by Arthur Kaser. ModeK cells were cultured in DMEM (DMEM Glutamax plus 10% FCS, non-essential amino acids and HEPES, Gibco). Cells were incubated at 37°C at 5% CO_2. All cell lines were authenticated by microscopic morphologic evaluation, characteristic growth curves and screening for mycoplasma.

Bacteria

The bacterial strains used in this study are listed on the Key Resources Table. *B. subtilis* (DSMZ, #10) and *E. coli*, strain K12 (DSMZ, #498) were cultured overnight at 37°C in LB medium. After adjustment to the desired density, serial dilutions of the inocula were plated on LB agar plates to verify the CFU.

METHOD DETAILS

Nested Sampling Design

The high microbial abundance (HMA) sponge species *Petrosia ficiformis*, *Chondrosia reniformis*, and *Agelas oroides* (each n=4) were collected at the Mongrí Coast, Cala Foradada, 3°12'00.09''E, 42°04'56.97''N, Girona, Catalunya, Spain) close to Barcelona by snorkelling and scuba diving within a 20 m radius. We randomly sampled four individuals per species. *Aplysina aerophoba* (n=4) was collected at a different site 25 km south-west (see for metadata Table S3). Immediately after collection, sponge samples were rinsed in sterile artificial seawater, plunge frozen in liquid nitrogen, and stored at -80°C until further processing. Prior to sponge sampling, for each spot, 30 litres of seawater were collected from the sponge vicinity using cooled sterilised tanks. Viroplankton was enriched by $FeCl_3$ flocculation according to John et al. (2011). Briefly, the seawater was pre-filtered (Millipore, 0.22 μm,142mm, GPWP14250) and the virus fraction was incubated with $FeCl_3$ (2.9 mg/liter) for 1 hour. The virion-iron precipitates were then recovered on a filter (GE Polycarbonate Membrane filter, 1.0 μm, 142mm, K10CP14220) and resuspended in EDTA-ascorbate buffer (pH 6).

Sample Processing and Virome Sequencing

Deep frozen sponge individuals were dissected, separating the outer epithelial area (pinacoderm) from the inner mesohyl matrix. All samples, including seawater references, were then randomly shuffled for virus purification and DNA/RNA extraction to avoid batch effects during processing. Samples were thawed in preboiled ice-cold extraction buffer (artificial seawater with 10mM EDTA and 3% (w/v) PVPP) and were disintegrated using a blender on ice at 6,500 rpm (T25 digital ULTRA-TURRAX, IKA). Particle aggregation was reduced by vortexing the suspension 10 min on ice. Tissue debris, PVPP bound secondary metabolites and bacterial cells were removed by centrifugation (2x 4,600g; 30 min at 4°C; ThermoScientific Heraeus Multifuge 3SR). The cleared supernatant was filtered through a 0.45 μm filter (Conceição-Neto et al., 2015), and virions were pelleted using a Beckman SW-41-Ti swinging bucket rotor at 135,000 x g for 2 h. Virions were re-suspended in modified SM-buffer containing 0.01 M Na_2S overnight, purified by low speed centrifugation at 4,300g for 5 min, loaded onto a CsCl gradient (1.7/1.5/1.3/1.2/1.1) according Thurber et al. (2009) and separated at 135,000 x g for 2 h. Virion-containing layers (CsCl density 1.2-1.5) were retrieved using a syringe and confirmed for viral particles by epifluorescence and transmission electron microscopy. Notably, this purification method was shown to enrich for bacteriophages and negatively selects for groups of eukaryotic viruses (Thurber et al., 2009). The virions were diluted in SM-buffer, purified by low speed centrifugation as before and pelleted at 135,000 x g for 2 h. Upon overnight resuspension in Tris buffer, and removal of undissolved particles at 1000 g for 1 min, the supernatant was transferred to a fresh tube and incubated with benzonase for 2 h at 37°C to remove free nucleic acid contamination. Encapsulated viral DNA and RNA were extracted according to Thurber et al. (2009) and Lachnit et al. (2015). Viral nucleotides were randomly amplified using a modified version of the Whole Transcriptome Amplification Kit 2 (WTA2, Sigma Aldrich) as described in Conceição-Neto et al. (2015). This approach allows the capture of single-stranded and double-stranded DNA and RNA viruses with little amplification bias (as sequence reads represent viral genome copies). NexteraXT libraries were prepared and sequenced on a Hiseq2500 run with 2x250 bp paired end reads at IKMB Kiel (Data S1). From the same tissue as used for the viromes, V1V2+V3V4 of the 16S rRNA gene was amplified and sequenced as described in Thomas et al. (2016).

Metagenome Cross-Assembly and Curation

Illumina reads were quality trimmed and cleaned from adapters, primers and reads with Ns or an average Q-score below 15 using BBMap v37.75 (https://sourceforge.net/projects/bbmap/). This Q-score threshold was confirmed by trimming reads to either Q15 or Q25 and by comparing the assembly statistics for each library using QUAST v5.0.2. (Gurevich et al., 2013). This showed inferior Q25 assemblies compared to Q15 assemblies as indicated by fewer long contigs (> 5 kb; see Data S1). The reads were then assembled per library (n=36) and in random subsets of the total library pool (50x 0.01%, 12x 0.05%, 12x 0.10%) using metaSPAdes v3.11.1 with default parameters (https://github.com/MartinTJahn/Iter_assembly). This multistep assembly strategy was tested in pilot assemblies to improve the quality of the assemblies as described in more detail in Coutinho et al. (2017). Contigs from all assemblies were clustered using a custom script (https://github.com/kseniaarkhipova/RedRed) into populations with mummer3 (Kurtz et al., 2004) at $\geq$95% ANI across $\geq$80% of their lengths as inspired by (Roux et al., 2017). To filter for viral sequences and to remove remaining potential cellular contamination, population contigs were submitted to VirSorter 1.0.3 (using Virome Database and Virome decontamination options) and were additionally classified with the Contig Annotation Tool (CAT; https://github.com/dutilh/CAT). Contigs above 5 kb that were VirSorter classified and/or had superkingdom classification "Viruses" in CAT were used for downstream genome-centric analysis. For functional gene-centric analyses, we increased stringency against cellular sequence contamination by considering only contigs with at least two VirSorter hits for viral hallmark genes (i.e., "major capsid protein," "portal", "terminase large subunit," "spike", "tail," "virion formation" or "coat) or CAT "viral superfamily" annotation. In the next round of cellular decontamination, we screened against single-copy prokaryotic marker genes using Anvi'o v.2.1.1 workflow (Eren et al., 2015). A proportion of 6.19% (79 of 1276) of the contigs were hit by the single-copy prokaryotic marker database. Manual curation ensured that most of the hits were homologous to phage nucleotide replication machinery (DNA/RNA polymerases) and RecA, while there were no hits against any ribosomal RNA indicative of low remining contamination levels with cellular DNA/RNA (Roux et al., 2013). One contig with the ClpX C4-type zinc finger domain was removed from further analysis due to its unclear viral association. These in silico filtration steps ensured that no cellular signals should have been included in the functional analysis.

Gene Content-Based Viral Clustering

Evolutionary relationships between the viral genome (fragments) were inferred by implementing reticulate classification based on gene sharing as developed by Lima-Mendez et al. (2008). Briefly, we predicted 869,624 proteins from a set of 46,307 viral sequences (see details below) with PRODIGAL v2.6.3 (Hyatt et al., 2010) and detected pairwise similarities using all-by-all BLASTp, requiring a minimal bit score 50. Protein families were identified with the Markov cluster algorithm (MCL) using inflation factor 2 (Enright et al., 2002). All viral genomes were then compared to each other for shared protein family content, and the probability that similarity was by chance was estimated using a hypergeometric formula (Lima-Mendez et al., 2008). The resulting significance scores were corrected for multiple comparisons, and genome pairs with scores $\geq$ 0 were joined by an edge [see https://github.com/kseniaarkhipova/LMCLUST]. To define viral clusters (VCs) in the genome network, we determined 1.4 as the best MCL inflation factor based on ICCC (intracluster clustering coefficient) maximization as described in Roux et al. (2015). The curated viral contigs were clustered with well-characterized isolate genomes downloaded from the Actinobacteriophage database project (http://phagesdb.org/; January 2018) and ViralRefseq (January 2018). To investigate overlap with other marine environments, we also clustered with viral sequences from 130 environmental virome libraries. Specifically, 78 viromes cross-assembled from seawater (incl. Tara Oceans), corals and sediment (Coutinho et al., 2017), 24 viromes from a seawater transect throughout the Mediterranean Sea (López-Pérez et al., 2017) and all viromes from sponges known to date supplemented with corals (Laffy et al., 2018). Viral clusters were taxonomically classified based on the placement of ViralRefseq entries in the network using a custom script [https://github.com/MartinTJahn/cluster_screener]. For each taxonomic rank, clusters were screened for Viral RefSeq entries and classified according to the supermajority (3/4) of their taxonomic annotations.

Abundance Profiles

Relative abundance patterns of viral genomes in the different samples were assessed by mapping quality control reads from each library against the curated genome-centric catalogue using BBMap 37.75 (option ambiguous=random, ANI $\geq$ 99%). The resulting 36 x 4484 count matrix was normalised for contig/virus length and library size to yield counts per kbp (CpK, Equation 1):

$$CpK = \frac{count}{length} * scaling\ factor * 10^3 \quad with\ scaling\ factor = \left(mean\ library\ size / sample\ library\ size \right) \qquad \text{(Equation 1)}$$

Community and Prevalence Classification

Distances between viral metagenomes were computed with the reference-independent cross-assembly (crAss;(Dutilh et al., 2012)) tool. The resulting clustering was calculated based on the SHOT formula and was drawn with iTOL (Letunic and Bork, 2019). To infer significant clustering of sample categories (type, species, tissue), the leaf labels and associated sample categories were compared to N=1,000 trees where the leaf labels were randomized using a custom script [treestats.pl; https://github.com/linsalrob/crAssphage/tree/master/bin]. The observed patterns were validated independently by hierarchical clustering based on Bray-Curtis distances of community abundance signatures calculated in the R package vegan (Dixon, 2003).

Viral population enrichment for sample types was assessed using the population enrichment score (Equation 2)

$$Pop.enrichment\ score = \frac{mean(CpK_{sample})}{mean(CpK_{other\ samples})} \qquad \text{(Equation 2)}$$

with a $\geq$ 2-fold enrichment considered to be enriched.

Viral population contigs (BCvir) were defined as detected in a sample when at least 75% of its length was covered by read mapping as suggested by Roux et al. (2017). BCvir were considered prevalent based on the supermajority rule when detected in at least 75% of the samples in a sample category. ''Individualists'' were those BCvir that were detected in only one individual but both tissues. ''Generalists'' were prevalent in all or several sample categories, while ''Specialists'' were prevalent in only one sample type (Data S1).

Annotation and Auxiliary Gene Classification

Proteins were predicted from viral contigs that passed the stringent cellular contamination filter (n=1,275 BCvir contigs) and were searched against the PFAM database (v31) using InterProScan v5.27-66.0 (Jones et al., 2014). Identified PFAM domains were then classified into 8 functional categories: ''metabolism'', ''lysis'', ''structural'', ''membrane transport, membrane-associated'', ''DNA replication, recombination, repair, nucleotide metabolism'', ''transcription, translation, protein synthesis'', ''other'', and ''unknown'' as in (Hurwitz et al., 2015) and extended by (Roux et al., 2016). This PFAM classification catalogue was augmented with manual classification of 85 PFAM signatures that were novel compared to the seawater viromes in the present study. We realized that in addition to auxiliary functions involved in the hosts metabolism (AMG), further categories might be relevant in the tripartite system of phage-prokaryote-eukaryote (PPE-interaction, hereafter). Therefore, we manually reannotated category ''others'' into classes ''signalling and protein-protein interaction'', ''cellular binding'' and ''cellular defence systems'', based on the literature research and functional evidence. The final extended PFAM classification catalogue for phages is given in Data S1 and is open for further use in other systems. For abundance estimations, multiple identical PFAM motifs in one protein, such as by repeats, were counted as one to ensure that quantification is not biased towards repeat domains. Sequences were also annotated with Prokaryotic Virus Orthologous Groups (pVOGs; (Grazziotin et al., 2017)) through HMMER 3.1b2 (hmmscan -E 10^{-5}), SEED subsystems through MG-RAST (E 10^{-5}) and were matched to NCBI-nr database entries using Diamond (e-value 10^{-5} and identity $\geq$40%). Proteins

were screened for terminal signal peptides using SignalP v4.1f and for transmembrane domains using TMHMM 2.0c. The tertiary structure of phage ankyrin-containing proteins was approximated by I-TASSER (default settings) (Roy et al., 2010). All annotations are combined in Data S1.

ANKp Expression and Purification

The 568 nt ANKp encoding phage gene (BCvir 4986 ORF 10) was optimized for *E. coli* codon usage and then de-novo synthesized in collaboration with GenScript (Piscataway, NJ, USA). The sequence was then cloned into the pET-22b(+) vector using BamHi and XhoI restriction sites. The same pET22b construct encoding GFP (Addgene plasmid # 38257 pET22b-GFP-CPDSalI) was used as a negative control and was a gift from Matthew Bogyo & Aimee Shen (Shen et al., 2009). Plasmids were transformed into BL21(DE3)-competent *E. coli*. Heterologous protein expression was induced in *E. coli* with 0.4 mM IPTG, and cultures were grown for 3 h at 37°C. For the native purification of the His-tagged target proteins, the Ni-NTA Fast Start Kit (Qiagen) was applied according to the manufacturer's instructions.

ANKp Cell Exposure Assays

E. coli K12 (DSMZ, #498) or *B. subtilis* (DSMZ, #10) overnight Luria–Bertani (LB) medium cultures were harvested at 4,000g for 10 min at room temperature and were washed twice with PBS. The bacterial suspension was then either incubated with ANKp (0 nM, 100 nM, 1 μM purified protein) or GFP (0 nM, 100 nM, 1 μM purified protein) in PBS for 10 min at 4°C under mild agitation. Prior to infection, the medium of BMDM cell culture was replaced with fresh antibiotic-free BMDM medium, and the cells were then infected with the pre-incubated *E. coli* K12 with protein in their medium using a range of multiplicities of infection (MOIs). Optical density measurements at a wavelength of 600 nm were performed using a Tecan Infinite 200 plate reader in a 96-well plate as described in Erez et al. (2017). In addition to assays with purified protein we performed the BMDM experiment with recombinant *E.coli* (see ANKp expression) directly expressing the proteins. As positive control for immune suppression we added the NF-κB inhibitor JSH-23 (Sigma-Aldrich). All *in vitro* data is representative for three independent non-randomized experiments. Sample sizes varied between the experimental approach used and were selected based on previous experience about the expected magnitude and variance of the phenotype. If not otherwise stated in the figure legends, all experiments included at least 3 biological replicates.

RNA Extraction and Quantitative RealTime PCR

Total RNA was isolated from BMDM and ModeK cell culture 16 h post infection using the RNeasy kit (Qiagen) and reverse transcribed using the Maxima H Minus First Strand cDNA Synthesis kit (Thermo Scientific). Quantitative RealTime PCRs were performed with TaqMan Gene Expression Master Mix (Applied Biosystems) according to the manufacturer's instructions and were analysed on the 7900HT Fast Real Time PCR System (Applied Biosystems). The applied TaqMan assays for pro-inflammatory markers are: *Cxcl1* (TaqMan ID Mm00433859_m1), *Gapdh* (TaqMan ID Mm99999915_g1), *Ifnb1* (TaqMan ID Mm00439552_s1), *Tnfα* (TaqMan ID Mm00443258_m1).

NF-κB–Dependent Luciferase Assay

The dual-luciferase assay using an NF-κB–dependent firefly luciferase (pNF-κB-Luc; Clontech) and a Renilla luciferase driven by the thymidine kinase promoter (pRLTK; Clontech) was performed according to the manufacturer's instructions. Briefly, ModeK cells cultured in DMEM (DMEM Glutamax plus 10% FCS, non-essential amino acids and HEPES, Gibco) were transfected with 20 ng pNF-κB-Luc and 3 ng pRL-TK using FuGENE 6 (Roche). Transfected cells were incubated for 24 h (37°C, 5% CO2), lysed and the lysate was subjected to the dual-luciferase assay carried out on a Tecan 96-well microplate reader.

QUANTIFICATION AND STATISTICAL ANALYSIS

Sample Sizes

n represents the number of sponge individuals, seawater replicates or cell assay experiments as described in legends of each figure.

Statistical Analysis

Topology of the crAss clustering (Figure 2) was compared against 1,000 random trees as detailed in the methods section. The reported consistency values represent the average of frequencies of the most frequent metadata annotation in a branch (Edwards et al., 2019). Cell assay data are presented as the mean ± SEM of at least three independent experiments. Statistical significance between treatments was determined by two-tailed unpaired Student's t-tests with p values less than 0.05 that were considered statistically significant. Statistical tests were performed using GraphPad Prism 6.0 Software.

DATA AND CODE AVAILABILITY

The accession number for all sequencing libraries, the cross-assembly, of both the sponge viromes and seawater references, as well as microbial amplicon data reported in this paper is GenBank: BioProject: PRJNA522695). All custom code is available at GitHub as indicated in the methods section.