

8 | Virology | Full-Length Text

# Sputnik virophage disrupts the transcriptional regulation of its host giant virus

Jingjie Chen,<sup>1</sup> Hiroyuki Ogata,<sup>1</sup> Hiroyuki Hikida<sup>1</sup>

**AUTHOR AFFILIATION** See affiliation list on p. 15.

ABSTRACT Sputnik virophages are small double-stranded DNA (dsDNA) viruses that replicate only inside host amoebae infected with giant dsDNA viruses, mimiviruses. Sputnik infection affects mimivirus replication, but their molecular interaction remains poorly understood. Here, we performed a time-course transcriptome analysis of Acanthamoeba castellanii cells infected with Acanthamoeba polyphaga mimivirus (APMV; hereafter referred to as Sputnik<sup>-</sup> cells) and those infected with both APMV and Sputnik 3 virophage (Sputnik<sup>+</sup> cells). The gene expression patterns of the amoeba were similar between these two conditions, whereas the expression of APMV genes was drastically affected by Sputnik, depending on the timing of their expression. Early-expressed APMV genes showed similar expression patterns in both conditions at the early stage of infection. However, at later stages, their expression levels remained higher in Sputnik<sup>+</sup> cells than in Sputnik<sup>-</sup> cells, suggesting a prolongation of early gene expression by Sputnik. Late-expressed APMV genes showed lower expression at earlier stages in Sputnik<sup>+</sup> cells, but their expression levels reached or exceeded those in Sputnik<sup>-</sup> cells at later stages, indicating a delay in gene expression. Overall, our results demonstrated that Sputnik infection drastically alters the transcriptome of APMV rather than amoeba, likely by disturbing the transition from early to late stages of APMV infection.

**IMPORTANCE** Virophages are small double-stranded DNA (dsDNA) viruses parasitizing other dsDNA viruses, such as giant viruses. Virophages inhibit the replication of giant viruses, and some protists use virophages as a defense system against giant viruses. However, molecular interactions among host cellular organisms, giant viruses, and virophages are largely unknown. Here, we performed a time-course transcriptome analysis of *Acanthamoeba castellanii* cells infected with Acanthamoeba polyphaga mimivirus (APMV) and those infected with both APMV and Sputnik 3 virophage. We demonstrated that the virophage has little effect on the amoeba transcriptome and primarily hijacks the transcriptional machinery of the giant virus. Furthermore, virophage infection alters giant virus gene expression, depending on their expression timing. The expression of early genes was prolonged, while that of late genes was delayed, suggesting that virophage infection disrupts the transition from the early to late stages of giant virus infection. This study provides molecular insights into the interactions within this unique tripartite system.

**KEYWORDS** virophage, giant virus, transcriptome analysis, eukaryotic virus, viruses of microorganisms, virus-virus interaction, gene regulation

V irophages are double-stranded DNA (dsDNA) viruses associated with other dsDNA viruses, nucleocytoviruses, which have large genomes and particles and are often called giant viruses (1–3). The first virophage, Sputnik, has an 18 kbp genome in 80 nm particles and is associated with a giant virus, mimivirus, which has a 1.2 Mbp genome in 700 nm particles. Sputnik and mimiviruses infect free-living amoebae of the genus

**Editor** Kristin N. Parent, Michigan State University, East Lansing, Michigan, USA

Address correspondence to Hiroyuki Hikida, hhikida@scl.kyoto-u.ac.jp.

The authors declare no conflict of interest.

See the funding table on p. 15.

Received 31 January 2025 Accepted 20 February 2025 Published 12 March 2025

Copyright © 2025 Chen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license. *Acanthamoeba*, but Sputnik cannot replicate without mimivirus infection. Furthermore, Sputnik infection reduces the infectivity of progeny mimiviruses, leading to the concept of "a virus parasitizing another virus" and the designation of virophage (4).

Following the discovery of Sputnik, related viruses have been isolated with mimiviruses and other nucleocytoviruses (1, 5), such as Mavirus associated with Cafeteria roenbergensis virus (CroV) (6) and Gezel-14T associated with Phaeocystis globosa virus-14T (PgV-14T) (7). In addition to the virophages isolated with these giant viruses, sequences of virophage relatives have been identified in various environmental metagenomic data (8–10), ranging from aquatic environments (11) to the animal gut rumen (12).

Not only Sputnik but also multiple isolated virophages also affect host nucleocytovirus replication by reducing progeny production (7, 13–17). However, the interaction between virophages and host nucleocytoviruses varies depending on the combinations. Sputnik virophages are thought to attach to the surface fibrils of host mimiviruses and be incorporated into host amoeba cells together (18). In contrast, Mavirus, a virophage of CroV, can enter host protist cells independently and wait for the corresponding giant virus infection (6, 17). Among the virophages parasitizing mimiviruses, Zamilon virophages replicate without affecting host-virus replication (16), unlike Sputnik virophages.

The expanded number of isolates and genomic sequences discovered from metagenomic data has provided novel insights into the vast diversity of this complex tripartite system. However, the molecular mechanisms underlying these microbial eukaryote-giant virus-virophage interactions remain largely unknown. In the present study, we focused on one of the tripartite systems, *Acanthamoeba castellanii*, Acanthamoeba polyphaga mimivirus (APMV), and Sputnik 3 virophage, analyzing their transcriptome dynamics. Our results strengthen the concept of virophages as viruses parasitizing other viruses and identify the disturbance of APMV transcriptional regulation caused by Sputnik infection.

# RESULTS

#### Transcriptome landscape of the tripartite system

Since the infection cycle of Sputnik was not fully explored, we first determined the time points for RNA sequencing based on the expression timing of marker genes: the DNA polymerase B (*polB*) gene of APMV and the major capsid protein (*mcp*) genes of both APMV and Sputnik (Fig. S1). The expression of APMV *polB* peaked at 3 hours post-infection (hpi), while the expression of APMV *mcp* peaked at 6 hpi and was maintained until 9 hpi. Sputnik *mcp* showed a similar pattern to APMV *mcp*. At 12 hpi, *polB* expression increased again, and *mcp* expression decreased, suggesting that the first infection cycle ends and secondary infection starts at this point. Based on these results, we performed transcriptome analysis at 0, 3, 6, and 9 hpi using amoeba cells infected only with APMV (Sputnik<sup>-</sup> cells) and those infected with both APMV and Sputnik (Sputnik<sup>+</sup> cells; Fig. 1).

In both conditions, the proportion of amoeba-derived reads decreased, while the reads from viruses increased throughout the infection (Fig. 2). In Sputnik<sup>-</sup> cells, the proportion of APMV-derived reads reached approximately 50% at 9 hpi, whereas in Sputnik<sup>+</sup> cells, the APMV and Sputnik reads each accounted for approximately 20% of the total reads at 9 hpi. The total proportion of viral reads in Sputnik<sup>+</sup> cells was around 40%, similar to the viral-read proportion in Sputnik<sup>-</sup> cells.

# Sputnik genes are divided into two groups by their expression timing

As very few reads were mapped to Sputnik genes at 0 hpi (Table S1), we considered that the Sputnik gene expression at 0 hpi was likely to be an artifact. Therefore, we investigated Sputnik gene expression at 3, 6, and 9 hpi. Sputnik genes were clustered into two groups based on their expression patterns (Fig. 3). The first group, early genes (five genes), exhibited high expression at 3 hpi, gradually decreasing from 6 to 9 hpi. The

Journal of Virology



FIG 1 A schematic image of the experiment. Cells and viruses were incubated for 1 hour, and the supernatant was exchanged for fresh medium to remove the remaining viruses. The time when the medium was exchanged was set as 0 hpi.

second group, late genes (17 genes), had low expression at 3 hpi, but their expression increased at 6 hpi and remained at a similar level until 9 hpi (Fig. 3a). Early genes included a gene annotated with a DNA replication function, while late genes included virion-related genes such as a DNA packaging protein, a membrane protein, and capsid proteins (Table S1).

# Sputnik affects a late stage of APMV infection

Principal component analysis (PCA) based on APMV gene expression showed that samples from Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells were closely located at 0 and 3 hpi (Fig. 4a). At these two time points, 0 and 130 APMV genes were differentially expressed, respectively (Fig. 4b). These results suggest similar transcription profiles of APMV at the early stages of infection between Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. However, at 6 and 9 hpi, samples from Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells were differentially expressed at these later time points, respectively (Fig. 4b). Moreover, samples from Sputnik<sup>+</sup> cells at 9 hpi overlapped with those from Sputnik<sup>-</sup> cells at 6 hpi in the PCA (Fig. 4a), suggesting that Sputnik infection inhibits the progression of APMV gene expression.



FIG 2 Proportions of viral and host reads at different time points. The proportion of reads from different sources relative to total reads (a) Sputnik<sup>-</sup> cells, where amoeba cells were infected only with APMV. (b) Sputnik<sup>+</sup> cells, where amoeba cells were infected with both APMV and Sputnik. R1, R2, and R3 represent each replicate.



**FIG 3** Gene expression pattern of Sputnik. (a) Heatmap and clustering based on normalized Sputnik gene expression. Gene expression was calculated as transcripts per million (TPM), followed by log2 transformation and normalized within each gene. In the left color bar, green and purple indicate early and late genes, respectively. (b) Principal component analysis (PCA) of Sputnik gene expression. Each circle represents a replicate, and each color represents a time point (n = 3).

# Sputnik infection has little effect on amoeba gene expression

In PCA based on amoeba gene expression, samples from Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells at each time point were closely located (Fig. 5a). No amoeba genes were differentially expressed at 0 and 3 hpi, while 90 (0.6%) and 98 (0.7%) amoeba genes were differentially expressed at 6 and 9 hpi, respectively (Fig. 5b; Table S2). Compared with the number of amoeba genes whose expression levels were altered by APMV infection (Fig. S2), these numbers were relatively small. Thus, we concluded that Sputnik infection has little effect on amoeba gene expression.

# Sputnik delays the expression of APMV genes

To elucidate the effect of Sputnik infection on APMV gene expression, we classified APMV genes based on their expression patterns in Sputnik<sup>-</sup> cells. To minimize noise from genes showing minor expression changes, we focused on differentially expressed genes (DEGs) for downstream analysis (see Materials and Methods). We identified four clusters: immediate-early, early, intermediate, and late (Fig. 6a; Table S3). These clusters mostly correspond to previous classifications (19). Immediate-early genes showed high expression at 0 hpi, with a gradual decline thereafter. Early genes peaked at 3 hpi, then decreased at subsequent time points. These genes mainly include those involved in host-virus interactions and DNA replication (Fig. S3; Table S3). Intermediate genes increased in expression from 0 to 3 hpi, maintained high levels at 6 hpi, and decreased



**FIG 4** Comparison of APMV gene expression level between Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. (a) PCA of APMV gene expression. Colors represent each time point, with circles and triangles representing replicates from Sputnik<sup>-</sup> cells and Sputnik<sup>+</sup> cells, respectively. (b) The expression changes of APMV genes are shown in volcano plots. APMV gene expression was compared between Sputnik<sup>-</sup> cells (as the baseline) and Sputnik<sup>+</sup> cells at each time point. Time points are indicated at the top left. Genes with an adjusted *P*-value ( $P_{adj}$ ) ≤0.05 and |log2 fold change| ≥ 1 are considered differentially expressed genes (DEGs) and are marked in red or blue. Red and blue indicate up-regulated and down-regulated genes, respectively. NS: no significant difference.

at 9 hpi. Late gene expression gradually increased from 6 to 9 hpi, and these genes are mainly involved in virion structure and morphogenesis (Fig. S3; Table S3).

We applied the same clustering method to the samples from Sputnik<sup>+</sup> cells, resulting in four analogous APMV gene clusters. While most genes clustered into the analogous clusters, some genes were grouped into clusters corresponding to later stages of expression (Fig. 6b; Table S3). This result aligns with the delay in APMV gene expression observed in the PCA (Fig. 4a).

# Prolonged gene expression in immediate-early and early genes

In APMV genes clustered into analogous groups, some genes exhibited noticeably different expression patterns between Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. For example, many intermediate genes showed relatively higher expression at 9 hpi in Sputnik<sup>+</sup> cells compared to Sputnik<sup>-</sup> cells (Fig. 6). To further examine APMV genes within each cluster, we performed subclustering based on the combined expression profiles in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells.



**FIG 5** Comparison of amoeba gene expression levels between Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. (a) PCA of amoeba gene expression. Circles and triangles represent each replicate from Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells, respectively. (b) The expression changes of amoeba genes are shown in volcano plots. Amoeba gene expression was compared between Sputnik<sup>-</sup> cells (as the baseline) and Sputnik<sup>+</sup> cells at each time point. Genes with  $P_{adj} \le 0.05$  and |log2 fold change|  $\ge 1$  are defined as differentially expressed genes (DEGs) and are marked in red or blue. Red and blue indicate up-regulated and down-regulated genes, respectively. NS: no significant difference.

Full-Length Text



**FIG 6** Clustering of APMV DEGs. Clustered heatmap of selected APMV DEGs from (a) Sputnik<sup>-</sup> and (b) Sputnik<sup>+</sup> cells. Gene expression was calculated as transcripts per million (TPM), followed by log2 transformation, and normalized within each time point. Rows and columns represent genes and time points, respectively. (a) The color bar indicates each cluster. (b) The color bar indicates the clusters identified in (a).

In the immediate-early genes, we identified four subgroups. However, subgroup 4 comprised two genes (Fig. 7a). Therefore, we focused on subgroups 1, 2, and 3. Genes in these subgroups exhibited comparable expression levels in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells at 0 and 3 hpi. A few genes in subgroup 1, a relatively small subgroup, showed slightly decreased expression in Sputnik<sup>+</sup> cells compared to Sputnik<sup>-</sup> cells at 6 and 9 hpi (Fig. 7b and c). In contrast, in subgroups 2 and 3, which include most immediate-early genes, the genes showed up-regulation at 6 hpi (Fig. 7b and c). This result suggests that



**FIG 7** Expression profile of APMV immediate-early genes. (a) Clustered heatmap showing normalized gene expression of immediate-early genes in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. Gene expression in both Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells was calculated as transcripts per million (TPM) and further normalized within each gene. The color bar represents the subgroups of genes. (b) Expression changes from Sputnik<sup>-</sup> to Sputnik<sup>+</sup> cells at each time point. The log2-fold change of each gene expression was calculated based on TPM. Each dot represents an APMV gene. (c) Temporal expression changes of representative genes. Genes with the highest expression in subgroups 1, 2, and 3 are shown. Error bars indicate the SD (n = 3).

Sputnik infection leads to prolonged expression of these genes. At 9 hpi, this prolonged expression ceased for the genes in subgroup 2 but continued for those in subgroup 3.

Early genes were divided into two subgroups (Fig. 8a). Genes in both subgroups showed similar temporal expression changes. Their expression increased from 0 to 3 hpi, with similar levels in both Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. However, at 6 and 9 hpi, these genes exhibited four- to eightfold higher expression in Sputnik<sup>+</sup> cells compared to Sputnik<sup>-</sup> cells (Fig. 8b and c). These results suggest that Sputnik infection also prolongs the expression of APMV early genes, similar to the effect observed with a part of immediate-early genes.

#### Delayed gene expression in intermediate and late genes

Intermediate genes were clustered into three subgroups (Fig. 9a). Sputnik infection resulted in slightly lower expression of genes in all three subgroups at 3 hpi, but their expression reached similar or higher levels than in Sputnik<sup>-</sup> cells at later stages of infection, suggesting a slight delay in the initiation of gene expression (Fig. 9b and c). Genes in subgroups 1 and 2 exhibited approximately fourfold higher expression in Sputnik<sup>+</sup> cells compared to Sputnik<sup>-</sup> cells at 6 hpi (Fig. 9b). Part of the genes in subgroup 3 showed lower expression in Sputnik<sup>+</sup> cells than in Sputnik<sup>-</sup> cells. However, the other genes in this subgroup also showed up-regulation, although the fold change was smaller than in the other subgroups. While the expression of the intermediate gene decreased at 9 hpi, the expression levels in Sputnik<sup>+</sup> cells were maintained higher than that in Sputnik<sup>-</sup> cells, increasing the gap in expression differences between the two conditions (Fig. 9b and c). These results suggest that intermediate genes also exhibit prolonged gene expression at late stages of infection.



**FIG 8** Expression profile of APMV early genes. (a) Clustered heatmap showing normalized gene expression of APMV early genes in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. Gene expression was calculated as transcripts per million (TPM), then normalized, and clustered as described in Fig. 7. (b) Expression changes from Sputnik<sup>-</sup> to Sputnik<sup>+</sup> cells at each time point. Each dot represents an APMV gene. (c) Temporal expression changes of representative genes. Genes with the highest expression from each subgroup are shown. Error bars indicate the SD (n = 3).

Late genes were clustered into two subgroups exhibiting similar expression patterns (Fig. 10a). At 0 hpi, genes in both subgroups showed low expression in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. Most genes showed lower expression in Sputnik<sup>+</sup> cells compared to Sputnik<sup>-</sup> cells at 3 and 6 hpi (Fig. 10b and c). This reduced expression was partially restored at 9 hpi, with some genes showing comparable expression levels in both conditions or even higher expression in Sputnik<sup>+</sup> cells than in Sputnik<sup>-</sup> cells (Fig. 10b). These results suggest that Sputnik infection also delays the expression timing of APMV late genes.

# Expression of APMV transcription-related genes precedes massive onset of Sputnik transcription

As mimiviruses encode complex transcriptional machinery, we examined the expression of 34 APMV genes with transcription-related annotations (Table S4). Based on the expression patterns, these transcription-related genes were categorized into four groups by clustering (Fig. 11a). Most transcription-related genes were classified as early and intermediate genes and already expressed at 3 hpi, preceding the massive onset of Sputnik transcription (i.e., late genes; Fig. 2 and 11). These genes showed higher expression in Sputnik<sup>+</sup> cells at the late stage of infection (i.e., 6 and 9 hpi), like other early and intermediate genes with prolonged expression. We also identified genes (i.e., putative cytidine deaminase and putative homeobox protein) that were classified as



**FIG 9** Expression profile of APMV intermediate genes. (a) Clustered heatmap showing normalized gene expression of APMV intermediate genes in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. Gene expression was calculated as transcripts per million (TPM), then further normalized, and clustered as described in Fig. 7. (b) Expression changes from Sputnik<sup>-</sup> to Sputnik<sup>+</sup> cells at each time point. Each dot represents an APMV gene. (c) Temporal expression changes of representative genes. Genes with the highest expression from each subgroup are shown. Error bars indicate the SD (n = 3).

immediate early genes, implying these genes are involved in the expression of the early Sputnik genes.

# Promoter analysis of APMV and Sputnik genes

A previous study (19) identified an early (AAAATTGA) and late (AT-rich element) promoter of APMV genes in the upstream region of some Sputnik genes. Among five early genes, we identified two genes with both early and late promoter motifs, two with late promoter, and one without any motif. Among 16 late Sputnik genes, we identified 8 genes with late promoter, 1 with early promoter, and 7 without any motif. These results suggest that Sputnik, at least in part, uses APMV promoters for its gene expression.

The shared promoter between APMV and Sputnik led to the hypothesis that Sputnik uses these promoters to exploit APMV transcription machinery (19). In this case, the expression of APMV genes with the promoter may be affected. We, therefore, investigated the expression level of APMV genes with the late promoter reported in the previous study (19). Out of the 95 APMV genes with the late promoter, 60 genes were differentially expressed in this study, indicating that Sputnik disrupts the expression of the genes with the late promoter. However, the expression patterns of these 60 genes were similar to those of the differentially expressed genes without the late promoter (Fig. 12). These results suggest that the use of the late promoter by Sputnik has little effect on the differential expression of APMV genes.

We further searched for new conserved motifs in the upstream regions of the APMV genes identified as DEGs. We identified the previously reported APMV early promoter motif AAAATTGA (19), its complemented motif, and a motif similar to the late promoter but failed to find other specific promoter motifs (Fig. S4). Also, a search for motifs in the two Sputnik gene clusters revealed no statistically significant motifs (Fig. S5).



**FIG 10** Expression profile for APMV late genes. (a) Clustered heatmap showing normalized gene expression of APMV late genes in Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells. Gene expression was calculated as transcripts per million (TPM), then further normalized, and clustered as described in Fig. 7. (b) Expression changes from Sputnik<sup>-</sup> to Sputnik<sup>+</sup> cells at each time point. Each dot represents an APMV gene. (c) Temporal expression changes of representative genes. Genes with the highest expression from each subgroup are shown. Error bars indicate the SD (n = 3).

# DISCUSSION

Sputnik virophages parasitize mimiviruses and reduce their propagation (4). A recent study reported the transcriptomic landscape of cells infected with another virophage, Zamilon, which has little effect on the propagation of host mimiviruses (20). Nevertheless, the molecular mechanisms underlying the interference by Sputnik in host mimivirus propagation are largely unknown. In the present study, we investigated the transcriptome landscape of amoeba cells infected with APMV and Sputnik. We found that Sputnik infection drastically alters the APMV gene expression pattern at the late stages of infection. The PCA of the APMV gene expression profile indicated that Sputnik infection inhibits the progression of mimivirus infection from 6 to 9 hpi. The expression patterns of late and intermediate genes further support this delay. These genes exhibited reduced expression at the early stages of infection, but their expression increased at the late stages in Sputnik-infected cells as found in cells only infected with APMV. Taken together, our results demonstrate that Sputnik infection disrupts the transcriptional regulation of APMV at the late stages of infection, similar to Zamilon infection.

Sputnik infection affects not only intermediate and late genes but also immediateearly and early genes. The expression of these genes typically starts at the early stages of infection and generally decreases at the late stages. Although these genes also showed decreased expression at the late stages in Sputnik-infected cells, the decrease was smaller compared to cells without Sputnik infection. Consequently, higher expression level was maintained at the late stages of infection in Sputnik-infected cells compared to those only infected with APMV. These results indicate that Sputnik infection prolongs



FIG 11 Expression of APMV transcription-related genes. (a) Clustered heatmap showing normalized gene expression of Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells for APMV transcription-related genes. Gene expressions were calculated as transcripts per million (TPM) and further normalized and clustered as described in Fig. 6. R1, R2, and R3 refer to each replicate. (b) Expression changes from Sputnik<sup>-</sup> and Sputnik<sup>+</sup> cells at each time point. Each dot represents an APMV gene. Colors represent subgroups.

the expression of APMV immediate-early and early genes. Meanwhile, this prolongation suggests that Sputnik infection may hinder the transition from the early to the late stages of APMV infection, disrupting APMV late gene expression (Fig. 13).

APMV immediate-early and early genes typically include those involved in DNA replication. Prolonged expression of these genes might benefit Sputnik, as it possesses a limited number of genes related to DNA replication. Also, transcription-related genes encoded by APMV showed similar prolonged expression, which may also benefit Sputnik gene expression. However, since most APMV genes have not been fully characterized, further studies are needed to understand the advantages of Sputnik in disturbing APMV gene expression.

Virophage was named after bacteriophage because its lifecycle depends on a host giant virus as bacteriophages depend on a host bacterium (4). This study revealed that Sputnik gene expression begins following APMV gene expression, including APMV transcription-related genes. We also found that Sputnik infection has little effect on both the proportion of amoeba-derived reads and the amoeba gene expression pattern. Combined with the recent study using Zamilon virophage (20), our results indicate that mimiviruses are primary hosts of Sputnik, supporting the original concept of virophage.

A previous study found promoter motifs that are shared between Sputnik and APMV genes and indicated that using the same promoter is a potential mechanism for virophages to exploit the transcriptional machinery of mimiviruses (19). Our results suggest that Sputnik partially uses the APMV promoter motifs for its gene expression.



**FIG 12** Effects of the shared promoter motif on APMV gene expression. Expression changes from Sputnik<sup>-</sup> to Sputnik<sup>+</sup> cells at each time point. Blue and yellow indicate genes with and without the shared promoter motif, respectively. Each dot represents an APMV gene classified as DEG.

However, the Sputnik infection disrupts APMV gene expression independent of the presence of the promoter. We also searched for additional promoter motifs for APMV but failed to detect these motifs. These results suggest that the shared promoter between APMV and Sputnik is not the main mechanism underlying the transcriptional changes in the APMV genes. Instead, we found that expression changes caused by Sputnik varied even among genes expressed at a similar timing. Although we failed to identify candidate genes, Sputnik might hijack specific transcriptional machinery to get precise manipulation of APMV genes. Combined with recent findings from other model systems of host giant virus-virophage interactions (17, 20), future studies may provide insights into the molecular mechanisms underlying this complex tripartite system.

# MATERIALS AND METHODS

# **Cells and viruses**

*A. castellanii* (Douglas) Page, strain Neff (ATCC 30010), was maintained in peptone-yeast extract-glucose (PYG) medium at 28°C. APMV and Sputnik 3 (4) were served as proto-types of mimivirus and virophage, respectively.

# **Titration of viruses**

The titer of APMV was measured using the 50% tissue culture infectious dose ( $TCID_{50}$ ) method on a 96-well plate. The titer of Sputnik was determined using a modified



FIG 13 A hypothetical model of how Sputnik infection affects APMV gene expression. Each box represents a period of APMV gene expression.

TCID<sub>50</sub> method. Briefly, amoeba cells were infected with APMV at a multiplicity of infection (MOI) of 1 on a 96-well plate, followed by inoculation with a serially diluted Sputnik solution. Wells containing infectious Sputnik were initially inferred by observing a reduced cytopathic effect under light microscopy compared to wells containing only APMV. The presence of Sputnik was confirmed by quantitative PCR (gPCR) as follows: the entire volume of supernatant and remaining cells was collected from wells at the highest dilution level containing infectious Sputnik, as well as from one order higher and lower dilution levels. Viral DNA was extracted from these wells using the following method: the collected supernatant was centrifuged at 15,000 rpm (Thermo Scientific, Sorvall ST 8FR centrifuge) for 1 hour at 4°C. The pellets were resuspended in 45 µL of 50 mM sodium hydroxide solution and incubated at 95°C for 5 minutes. Then, 5 µL of 1 M Tris(hydroxymethyl)aminomethane hydrochloride (pH 8.0) and 450 µL Tris-EDTA buffer (10 mM Tris, 1 mM EDTA, and pH8.0) were added to the solution, which was then subjected to qPCR. The reaction mixture was prepared using the KAPA PROBE Fast gPCR kit (ROXLowgPCR, KAPA BIOSYTEMS), with 0.3 µM forward and reverse primers and 0.4 µM probe. The PCR conditions were as follows: 95°C for 3 min; 40 cycles of 95°C for 3 s, 60°C for 20 s, and 72°C for 1 s. The sequences of the primers and probe (21) are shown in Table S5.

# Sample preparation

A total of  $1 \times 10^6$  amoeba cells were inoculated with APMV or APMV and Sputnik at an MOI of 10 in PYG medium. The inoculated cells were incubated for 1 hour at room temperature with gentle shaking. After incubation, the culture medium was replaced

with 1 mL of phosphate-buffered saline (PBS) to remove uninfected viruses. The PBS was then replaced with 1.5 mL of fresh PYG medium, and the cells were incubated at 30°C for the designated periods. The time point at which the medium was exchanged for fresh PYG medium was designated as 0 hpi.

# **RNA** extraction

Infected cells were collected at 0, 1, 3, 6, 9, and 12 hpi using a cell scraper (Thermo Fisher Scientific) and then centrifuged at 2,000 rpm for 5 min at 4°C. The cell pellet was stored in 1 mL of TRIzol (Invitrogen) at  $-80^{\circ}$ C. Three independent experiments were conducted as biological replicates. RNA extraction followed the manufacturer's protocol. Briefly, 20% of the sample volume of chloroform was added to the cells in TRIzol. The RNA-containing aqueous layer was transferred to new tubes to which 500 µL of isopropanol was added to precipitate RNA. The RNA pellet was washed with 70% ethanol, dissolved in nuclease-free water, and treated with DNase I (New England Biolabs). RNA was then re-extracted using phenol/chloroform and re-precipitated with 99.5% ethanol and 0.3 M sodium acetate. The precipitate was washed with 70% ethanol and dissolved in nuclease-free water. RNA concentration was measured using a Qubit 4 Fluorometer (Invitrogen) with the Qubit RNA BR Assay Kit (Invitrogen). RNA purity was assessed by measuring the A260/A280 and A260/A230 ratios using a DeNovix DS-11 spectrophotometer (SCRUM Inc.).

#### **RNA** sequencing

Quality control and sequencing of RNA samples were performed by Rhelixa Inc. (Japan). In brief, strand-specific libraries were prepared using poly-A selection and sequenced to a depth of 5 G bases per sample using Illumina NovaSeq 6000.

#### Measuring gene expression by qPCR

Extracted RNA was adjusted to a uniform concentration across all samples using nuclease-free water. For cDNA synthesis, 0.5  $\mu$ g of total RNA was used with D(T)23VN primer (New England Biolabs), 10 mM dNTPs (New England Biolabs), AMV reverse transcriptase (New England Biolabs), and an RNase inhibitor (New England Biolabs). The reaction was conducted at 42°C for 1 hour, followed by enzyme deactivation at 80°C for 5 min.

The synthesized cDNA was then subjected to qPCR using the KAPA SYBR Fast qPCR kit (ROXLowqPCR, KAPA BIOSYTEMS) with 0.2  $\mu$ M of forward and reverse primers. The sequences of primers (21, 22) are listed in Table S5. The PCR conditions were as follows: 95°C for 20 s; 40 cycles of 95°C for 15 s, 60°C for 20 s, and 72°C for 10 s. Relative gene expression was calculated using the 2<sup> $\Delta$ Ct</sup> method, with the highest expression level set as 100.

#### Data analysis

The quality of reads was assessed using FastQC (v0.12.0) (23). With an overall quality score exceeding 20 and no known adaptors detected, no trimming was performed. Reads were mapped to a sequence data set consisting of the Sputnik 3 genome (NC\_011132.1) and the APMV genome (NC\_014649.1) using HISAT2 (v 2.2.1) (24), with a maximum intron size of 5,000 bp. Unmapped reads were subsequently mapped to the *A. castellanii* genome (GCF\_000313135.1\_Acastellanii.strNEFF\_v1) with a maximum intron size of 500,000 bp. Output data were processed with Samtools (v1.19) (25). The number of reads mapped to each gene was counted using HTSeq (v2.0.5) (26) in union mode with the reverse strand-specific assay option. Gene expression levels were individually normalized to transcripts per million (TPM) by using summed reads mapped to Sputnik, APMV, and amoeba genome.

PCA and DEG detection were conducted using DESeq2 (1.42.0) (27) with read count data for each gene. Genes were classified as DEGs with a false discovery rate <0.05 and absolute log2-fold change  $\geq 1$  at any time point (28).

DEGs were clustered using the k-means method. Log2-transformed TPM for each gene was normalized by adjusting the mean to 0 and the variance to 1 before clustering. The optimal number of clusters for Sputnik and APMV was determined using the total-within-sum-of-squares method in factoextra (v1.0.7).

Annotation information for APMV genes was retrieved from the National Center for Biotechnology Information GenBank database. Functional categories for each APMV gene were assigned manually, referring to previous studies (19, 29–31).

#### **Promoter search**

One hundred base pair sequences upstream of the open reading frames of each DEG in Fig. 5 matching the same expression timing were extracted. Sequence motifs were analyzed using MEME Suite 5.5.4 (32). Previously identified motifs were searched by FIMO using settings in a previous study (19). Also, sequence motifs were predicted *de novo* by MEME. MEME was run in classic mode with motif width ranges set from 8 to 25 bp and the "any number of repetitions" option.

#### Visualization

All data processing was carried out using the dplyr (v1.0.10) (33) and tidyr (v1.3.0) (34) packages in R (v4.2.1) (35). Heatmaps were generated with the pheatmap (v1.0.12) package (36). All plots were created using the ggplot2 (v3.4.4) package (37).

# ACKNOWLEDGMENTS

We thank Prof. Bernard La Scola and Ms. Lina Barrassi for kindly providing APMV and Sputnik 3 and Dr. Matthias Fischer and Dr. Anna Koslová for kindly sharing the method of virophage titration. Computation time was provided by the SuperComputer System, Institute for Chemical Research, Kyoto University.

This study was supported by the Japan Society for the Promotion of Science KAKENHI grant numbers 22K15175 to H.H. and 22H00384 to H.O. and by Japan Science and Technology Agency ACT-X Grant number JPMJAX22BI to H.H.

J.C., H.O., and H.H. conceptualized the study. J.C. performed experiments and formal analysis. J.C. and H.H. wrote the manuscript with intellectual input from all authors. H.H. and H.O. supervised the project.

# **AUTHOR AFFILIATION**

<sup>1</sup>Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Japan

# **AUTHOR ORCIDs**

Jingjie Chen D http://orcid.org/0009-0006-4574-5761 Hiroyuki Ogata D https://orcid.org/0000-0001-6594-377X Hiroyuki Hikida D http://orcid.org/0000-0002-4876-3133

# FUNDING

Funder	Grant(s)	Author(s)
MEXT   Japan Society for the Promotion of Science (JSPS)	22K15175	Hiroyuki Hikida
MEXT   Japan Society for the Promotion of Science (JSPS)	22H00384	Hiroyuki Ogata
MEXT   Japan Science and Technology Agency (JST)	JPMJAX22BI	Hiroyuki Hikida

#### **AUTHOR CONTRIBUTIONS**

Jingjie Chen, Conceptualization, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review and editing | Hiroyuki Ogata, Conceptualization, Funding acquisition, Supervision, Writing – review and editing | Hiroyuki Hikida, Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review and editing |

# DATA AVAILABILITY

The RNA sequencing data are available at NCBI/ENA/DDBJ under accession number PRJDB18584.

# **ADDITIONAL FILES**

The following material is available online.

#### Supplemental Material

Supplemental material (JVI00192-25-s0001.docx). Figures S1 to S5; Table S5. Additional supplemental tables (JVI00192-25-s0002.xlsx). Tables S1 to S4.

#### REFERENCES

- 1. Duponchel S, Fischer MG. 2019. Viva lavidaviruses! Five features of virophages that parasitize giant DNA viruses. PLOS Pathog 15:e1007592. https://doi.org/10.1371/journal.ppat.1007592
- Dance A. 2021. Beyond coronavirus: the virus discoveries transforming biology. Nature New Biol 595:22–25. https://doi.org/10.1038/d41586-02 1-01749-7
- Koonin EV, Senkevich TG, Dolja VV. 2006. The ancient Virus World and evolution of cells. Biol Direct 1:29. https://doi.org/10.1186/1745-6150-1-29
- La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, Raoult D. 2008. The virophage as a unique parasite of the giant mimivirus. Nature New Biol 455:100–104. https://doi.org/10.1038/nature07218
- Claverie J-M, Abergel C. 2009. Mimivirus and its virophage. Annu Rev Genet 43:49–66. https://doi.org/10.1146/annurev-genet-102108-134255
- Fischer MG, Suttle CA. 2011. A virophage at the origin of large DNA transposons. Science 332:231–234. https://doi.org/10.1126/science.1199 412
- Roitman S, Rozenberg A, Lavy T, Brussaard CPD, Kleifeld O, Béjà O. 2023. Isolation and infection cycle of a polinton-like virus virophage in an abundant marine alga. Nat Microbiol 8:332–346. https://doi.org/10.1038 /s41564-022-01305-7
- Paez-Espino D, Zhou J, Roux S, Nayfach S, Pavlopoulos GA, Schulz F, McMahon KD, Walsh D, Woyke T, Ivanova NN, Eloe-Fadrosh EA, Tringe SG, Kyrpides NC. 2019. Diversity, evolution, and classification of virophages uncovered through global metagenomics. Microbiome 7:157. https://doi.org/10.1186/s40168-019-0768-5
- Coutinho FH, Gregoracci GB, Walter JM, Thompson CC, Thompson FL. 2018. Metagenomics sheds light on the ecology of marine microbes and their viruses. Trends Microbiol 26:955–965. https://doi.org/10.1016/j.tim. 2018.05.015
- Zhou J, Zhang W, Yan S, Xiao J, Zhang Y, Li B, Pan Y, Wang Y. 2013. Diversity of virophages in metagenomic data sets. J Virol 87:4225–4236. https://doi.org/10.1128/JVI.03398-12
- Oh S, Yoo D, Liu W-T. 2016. Metagenomics reveals a novel virophage population in a tibetan mountain lake. Microbes Environ 31:173–177. htt ps://doi.org/10.1264/jsme2.ME16003
- Yutin N, Kapitonov VV, Koonin EV. 2015. A new family of hybrid virophages from an animal gut metagenome. Biol Direct 10:19. https://d oi.org/10.1186/s13062-015-0054-9
- Tokarz-Deptuła B, Czupryńska P, Poniewierska-Baran A, Deptuła W. 2018. Characteristics of virophages and giant viruses. Acta Biochim Pol 65:487–496. https://doi.org/10.18388/abp.2018\_2631
- 14. Mougari S, Chelkha N, Sahmi-Bounsiar D, Di Pinto F, Colson P, Abrahao J, La Scola B. 2020. A virophage cross-species infection through mutant

selection represses giant virus propagation, promoting host cell survival. Commun Biol 3:248. https://doi.org/10.1038/s42003-020-0970-9

- Mougari S, Bekliz M, Abrahao J, Di Pinto F, Levasseur A, La Scola B. 2019. Guarani virophage, a new sputnik-like isolate from a Brazilian Lake. Front Microbiol 10:1003. https://doi.org/10.3389/fmicb.2019.01003
- Gaia M, Benamar S, Boughalmi M, Pagnier I, Croce O, Colson P, Raoult D, La Scola B. 2014. Zamilon, a novel virophage with mimiviridae host specificity. PLoS ONE 9:e94923. https://doi.org/10.1371/journal.pone.009 4923
- Fischer MG, Hackl T. 2016. Host genome integration and giant virusinduced reactivation of the virophage mavirus. Nature New Biol 540:288–291. https://doi.org/10.1038/nature20593
- Boyer M, Azza S, Barrassi L, Klose T, Campocasso A, Pagnier I, Fournous G, Borg A, Robert C, Zhang X, Desnues C, Henrissat B, Rossmann MG, La Scola B, Raoult D. 2011. Mimivirus shows dramatic genome reduction after intraamoebal culture. Proc Natl Acad Sci U S A 108:10296–10301. ht tps://doi.org/10.1073/pnas.1101118108
- Legendre M, Audic S, Poirot O, Hingamp P, Seltzer V, Byrne D, Lartigue A, Lescot M, Bernadac A, Poulain J, Abergel C, Claverie J-M. 2010. mRNA deep sequencing reveals 75 new genes and a complex transcriptional landscape in mimivirus. Genome Res 20:664–674. https://doi.org/10.110 1/gr.102582.109
- Bessenay A, Bisio H, Belmudes L, Couté Y, Bertaux L, Claverie J-M, Abergel C, Jeudy S, Legendre M. 2024. Complex transcriptional regulations of a hyperparasitic quadripartite system in giant viruses infecting protists. Nat Commun 15:8608. https://doi.org/10.1038/s41467 -024-52906-1
- Gaia M, Pagnier I, Campocasso A, Fournous G, Raoult D, La Scola B. 2013. Broad spectrum of mimiviridae virophage allows its isolation using a mimivirus reporter. PLoS ONE 8:e61912. https://doi.org/10.1371/journal. pone.0061912
- Ngounga T, Pagnier I, Reteno D-GI, Raoult D, La Scola B, Colson P. 2013. Real-time PCR systems targeting giant viruses of amoebae and their virophages. Intervirology 56:413–423. https://doi.org/10.1159/00035456 3
- 23. Andrews S. 2010. A quality control tool for high throughput sequence data
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915. https://doi.org/10.1038/s41587-019-0201-4
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008. https://doi.org/10.10 93/gigascience/giab008

- Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F. 2022. Analysing highthroughput sequencing data in Python with HTSeq 2.0. Bioinformatics 38:2943–2945. https://doi.org/10.1093/bioinformatics/btac166
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. http s://doi.org/10.1186/s13059-014-0550-8
- Rodrigues RAL, Louazani AC, Picorelli A, Oliveira GP, Lobo FP, Colson P, La Scola B, Abrahão JS. 2020. Analysis of a marseillevirus transcriptome reveals temporal gene expression profile and host transcriptional shift. Front Microbiol 11:651. https://doi.org/10.3389/fmicb.2020.00651
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4:41. https://doi.org/10.1186/1471-2105-4-41
- Koonin EV, Yutin N. 2010. Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology 53:284–292. https://doi.o rg/10.1159/000312913

- Yutin N, Wolf YI, Raoult D, Koonin EV. 2009. Eukaryotic large nucleocytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol J 6:223. https://doi.org/10.1186/174 3-422X-6-223
- 32. Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. Nucleic Acids Res 43:W39–49. https://doi.org/10.1093/nar/gkv416
- Wickham H, François R, Henry L, Müller K, Vaughan D. 2023. \_dplyr: a grammar of data manipulation\_. R package version 1.1.4. Available from: https://CRAN.R-project.org/package=dplyr
- Wickham H, Vaughan D, Girlich M. 2024. \_tidyr: tidy messy data\_. R package version 1.3.1 https://CRAN.R-project.org/package=tidyr
- 35. R Core Team. 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Kolde R. 2019. \_pheatmap: pretty heatmaps\_. R package version https:// CRAN.R-project.org/package=pheatmap
- 37. Wickham H. 2016. Ggplot2: elegant graphics for data analysis. Springer-Verlag New York.