

RESEARCH ARTICLE

Detailed single-cell mapping of the transcriptional response to a virus infection driven by copy-back viral genomes

Yanling Yang, Emna Achouri, Munyaradzi Tambo, Carolina B. López^{*}

Department of Molecular Microbiology and Center For Women Infectious Disease Research, Washington University School of Medicine, St. Louis, Missouri, USA

* clopezzalaquett@wustl.edu



OPEN ACCESS

Citation: Yang Y, Achouri E, Tambo M, López CB (2026) Detailed single-cell mapping of the transcriptional response to a virus infection driven by copy-back viral genomes. PLoS Pathog 22(4): e1014184. <https://doi.org/10.1371/journal.ppat.1014184>

Editor: Makoto Takeda, University of Tokyo Graduate School of Medicine Faculty of Medicine: Tokyo Daigaku Daigakuin Igakukei Kenkyuka Igakubu, JAPAN

Received: January 20, 2026

Accepted: April 21, 2026

Published: April 27, 2026

Copyright: © 2026 Yang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: This work was supported by the National Institute of Allergy and Infectious

Abstract

The antiviral response to several clinically significant viruses, including respiratory syncytial virus and parainfluenza virus, is driven by copy-back viral genomes (cbVGs) generated during virus replication. However, the broader impact of cbVGs on the functional states of host cells remains undefined. Here, we developed a single-cell RNA-sequencing and computational framework to map cbVG-driven host responses during Sendai virus infection. Unsupervised profiling identified distinct transcriptional states throughout the course of infection, highlighting a shift from early antiviral signaling to later inflammatory and remodeling programs. Stratifying infected cells by cbVG status demonstrated that cbVG-positive cells initiate interferon and chemokine programs, which later spread to cbVG-negative cells. At later stages, cbVG-positive cells acquire additional signaling, cytoskeletal, transcriptional, and stress-adaptation programs, which are absent in cbVG-clean infection. This work defines the broader cbVG-driven layered and dynamic host response and provides a valuable high-resolution resource of the temporal cellular response to a virus infection.

Author summary

Viruses often produce non-standard viral genomes during replication. One important class of these molecules, called copy-back viral genomes (cbVGs), is known to strongly stimulate antiviral immune responses. However, it remains unclear how cbVGs influence the behavior of individual infected cells and how these responses spread across a population of cells during infection. In this study, we developed a single-cell sequencing approach that allows detection of cbVGs together with host gene expression in the same cell during Sendai virus infection. This approach enabled us to track how different groups of infected cells respond over time. We found that cells containing cbVGs initiate strong antiviral and interferon responses early in infection. These signals later spread

Diseases, National Institutes of Health (grants AI137062 and AI188900 to CBL), by a BJC Investigator Program award to CBL, by the Alexander & Gertrude Berg Fellowship awarded to YY, and by the Helen Hay Whitney Foundation Postdoctoral Fellowship awarded to MT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

to neighboring infected cells that do not contain detectable cbVGs, likely through secreted cytokines. At later stages of infection, cbVG-positive cells adopt additional stress and recovery programs that are not observed when cbVGs are absent. Together, our findings reveal how cbVGs shape the timing and diversity of antiviral responses at the single-cell level and provide a high-resolution view of how host cells dynamically respond to viral infection.

Introduction

Non-standard defective viral genomes (nsVGs) are generated in nearly all RNA virus infections and have long been recognized as potent activators of innate immunity [1–3]. Among them, copy-back viral genomes (cbVGs) emerge during negative-sense RNA virus replication and strongly induce the antiviral interferon response through triggering signaling by cytosolic RNA sensors [4–12]. Although cbVGs have been detected across multiple virus families and shown to impact infection dynamics and clinical outcomes [7, 13–25], how cbVGs shape host responses through the infection course remains unresolved. A fundamental barrier has been the lack of single cell data that allows building a conceptual model to explain how a rare cbVG-producing subpopulation of infected cells drives antiviral programs across an infected tissue. So far, studies on the impact of cbVGs in antiviral immunity have either relied on bulk measurements of infected populations [26–30], which obscure cell-to-cell heterogeneity and mask cell-state transitions, or have focused mostly on the expression of interferons [31, 32] overlooking the broader impact of cbVG-driven signals through time and across the infected population.

The biology of cbVGs has primarily been investigated after preparing cbVG-high virus stocks or adding purified cbVGs to infected cultured cells [33, 34], or after isolating cbVG-rich populations by flow cytometry following RNA-FISH [35]. While these approaches established cbVGs as strong inducers of antiviral signaling, they lack the resolution to distinguish cbVG-positive (cbVG⁺) cells from those responding indirectly (cbVG⁻) within the same population. Moreover, cbVG formation is stochastic, producing diverse species that vary in abundance across cells [14]. As a result, early signaling events initiated by a small cbVG⁺ subpopulation are diluted in bulk measurements, preventing analysis of temporal dynamics, paracrine effects, and cbVG-driven programs that emerge later in infection. Together, these limitations have left several gaps in the understanding of the early events of antiviral immunity, including understanding how cbVGs orchestrate population-wide responses, dissecting transcriptional programs directly driven by cbVGs versus secondary effects of antiviral signaling, and importantly, understanding the full impact of cbVGs on the infected cell.

Recent advances in single-cell RNA sequencing (scRNA-seq) enable simultaneous profiling of host and viral transcriptomes in thousands of individual cells [36–40]. Single-cell approaches have revealed heterogeneity in viral gene expression, innate immune activation, and cell-state transitions during infection [41–44]. However, direct detection of cbVGs at single-cell resolution remains technically challenging. CbVGs

can only be distinguished from the standard viral genome by recognizing the junction sequence that forms when the viral polymerase falls off the template strand during replication and rejoins the nascent RNA strand to continuing copying this strand of the genome. In addition, most RNA viruses produce populations of cbVG that differ in junction positions [14,17,18,22,45–47]. Standard scRNA-seq relies on identification of polyadenylated transcripts and captures only the end of transcripts near the template-switch oligo [48,49]. As cbVGs are not polyadenylated, and cbVG junction sequences vary in their position relative to the template-switch oligo, most cbVGs are lost during standard scRNA-seq library preparation.

Here, we took advantage of the murine parainfluenza virus Sendai strain Cantell (SeV C) to develop an integrated single-cell sequencing and computational framework to directly identify cbVGs in individual cells during infection. SeV C produces a single dominant cbVG (cbVG-546) that potently induces interferon expression [13]. Importantly, the junction sequence of cbVG-546 is detectable by the conventional 10x Genomics scRNA-Seq approach. Unsupervised profiling uncovered dynamic cellular states of SeV-positive cells, revealing cbVG-associated transitions from early interferon responses to later inflammatory programs. By resolving SeV-positive cells into cbVG⁺, cbVG⁻ across time, we define cbVG-driven transcriptional programs. Single-cell profiling at 6, 12, and 24 hours post infection revealed that cbVGs initiate interferon and chemokine programs from a rare cbVG⁺ subpopulation at early time points, which subsequently propagate to cbVG⁻ cells. At later time points, cbVG⁺ cells express additional transcriptional modules that remain poorly represented in cbVG⁻ cells, defining cbVG-specific programs beyond interferon responses. These results revealed a layered host response to virus infection driven by cbVGs, provide a single-cell roadmap for studying the impact of cbVGs in other infections, and serve as a resource for dissecting the components of the antiviral response within the first 24 h of infection.

Results

Single-cell sequencing enables direct detection of cbVGs in individual infected cells

cbVGs are generated by the viral RNA-dependent RNA polymerase during genome replication and contain a unique rejoin junction sequence that distinguishes them from standard viral genomes (S1 Fig). The VODKA2 pipeline detects cbVGs from short read RNA-seq datasets by identifying these junction-spanning reads (S1B Fig). To directly resolve cbVG-producing cells in single-cell data, we modified the standard scRNA-seq workflow by incorporating a virus-specific reverse transcription primer [50–52] in addition to the standard oligo(dT) priming for capturing mRNA (Fig 1A). The virus-specific primer was designed to capture viral transcripts including the junction sequence of cbVG-546, the dominant cbVG species generated during infection with SeV C [13]. Both primers share a common 5' overhang sequence, enabling recovery of host mRNAs, viral mRNAs, antigenomes, and cbVG transcripts during library preparation (Fig 1B). Junction-spanning reads from cbVG-546 can be captured because the transcript containing the junction is 278bp in length and the rejoin site is positioned 94bp from the template-switch oligo, within the read length of 10x-Illumina sequencing (Fig 1C). To characterize cbVG-driven responses over time, A549 cells were infected with SeV C at a multiplicity of infection (MOI) of 1.5–2 and sampled at 6, 12, and 24 hpi. An independent infection at higher MOI (MOI=5) was performed to increase the probability of detecting early cbVG-induced programs and to test reproducibility across infection settings. For each condition, two biological replicates were sequenced and subsequently combined for downstream analyses. Sequencing data were processed by integrating a standard scRNA-seq pipeline with VODKA2 (S2 Fig). Given the limited capture efficiency of viral transcripts in single-cell RNA-seq, cells were classified as SeV-positive (SeV⁺) or cbVG-positive (cbVG⁺) if at least one SeV or cbVG Unique Molecular Identifier (UMI) was detected. The number of cells passing quality control, as well as the number of SeV⁺ and cbVG⁺ cells per condition, are shown in Fig 1D. Although the number of infected cells was similar across conditions, the number of cbVG⁺ cells increased with time and MOI, consistent with progressive cbVG accumulation during infection. Viral UMI counts increased with MOI at 6 hpi, and with infection time at later time points, with 1–30 cbVG UMI per cbVG⁺ cell detected across samples (Fig 1E). Together, these results demonstrate that our modified scRNA-seq approach reliably detects viral and cbVG reads across infection time points at single-cell resolution, enabling direct identification of cbVG⁺ cells during SeV infection.

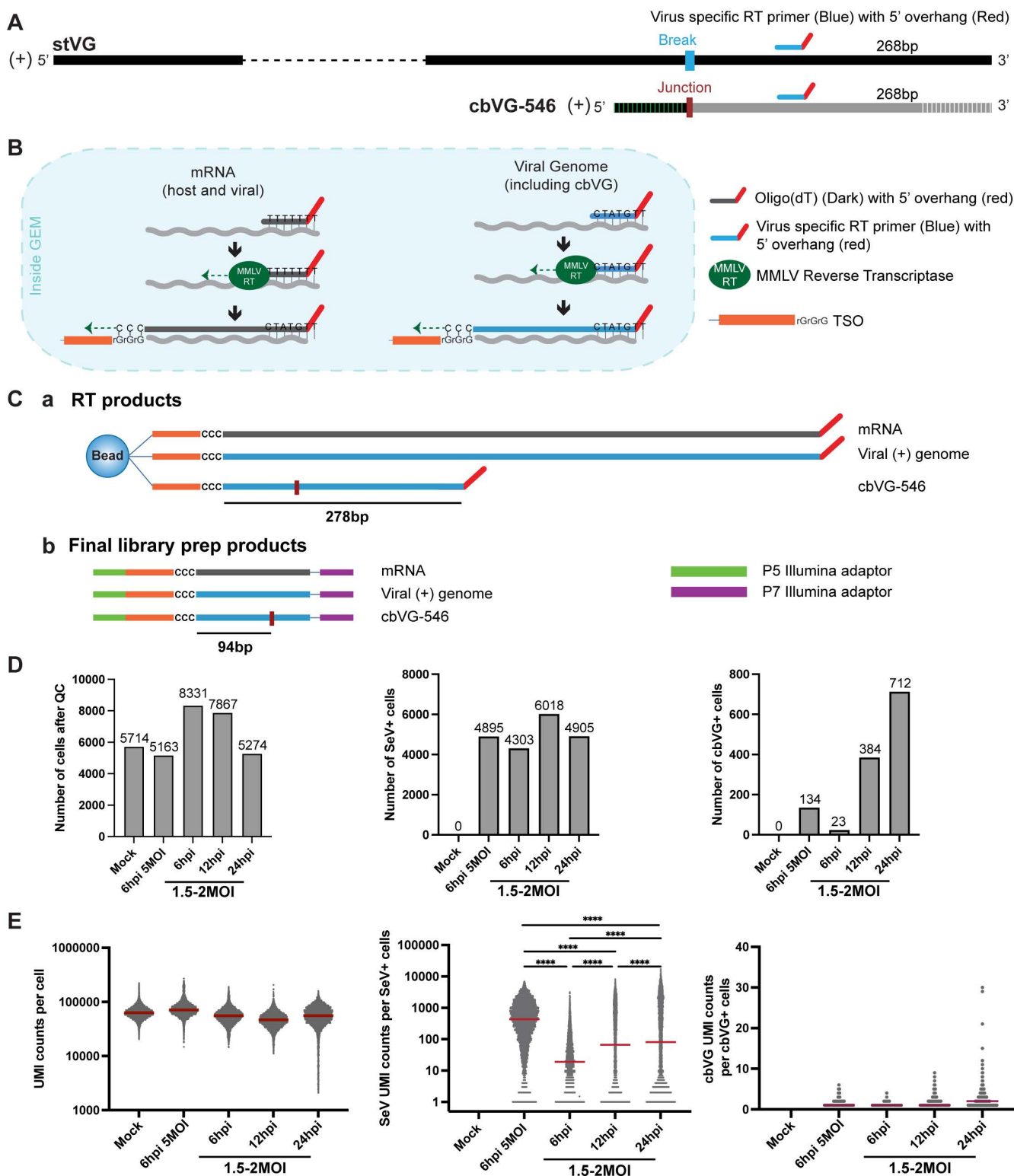


Fig 1. Detection of cbVGs at the single-cell level. (A) A virus-specific RT primer (blue) with a 5' overhang (red) was designed to bind 256 bp from the 3' end of the viral genome to capture both stVGs and cbVG-546. (B) Diagram showing mRNA and viral RNA capture and reverse transcription (RT) inside 10x Genomics Chromium droplets. (C) Illustration of RT and library preparation for Illumina sequencing. (a) RT products of mRNA, viral genome,

and cbVG-546 with the cbVG junction region labeled. A 278 bp of cbVG-546 that includes the junction region is transcribed. (b) Final library prep products with Illumina adaptors (green and purple) added to both ends. The cbVG-546 junction is 94 bp from the TSO. (D) Number of cells passing quality control (QC) and numbers of SeV-positive and cbVG-positive cells in each sample. (E) Total UMI counts, SeV reads, and cbVG reads per cell in each sample. Statistical differences in UMI counts, SeV reads or cbVG reads per cell across clusters were assessed using the Kruskal-Wallis test. Following a significant overall test, pairwise comparisons were performed using Dunn's post-hoc test. Significance levels are indicated as **** $p < 0.0001$.

<https://doi.org/10.1371/journal.ppat.1014184.g001>

Single-cell analysis identifies a rare cbVG⁺ subpopulation as the origin of early interferon induction

To define the relationship between the different types of viral genomes and early interferon responses, we performed unbiased clustering of infected cells based on their transcriptomes and overlaid viral RNA content, including both standard viral genomes and cbVG junction reads (Fig 2). While SeV RNA was distributed across all clusters, cbVGs were enriched in a subset of transcriptional states, indicating that cbVG production is not uniform across infected cells (Fig 2A–2B). A

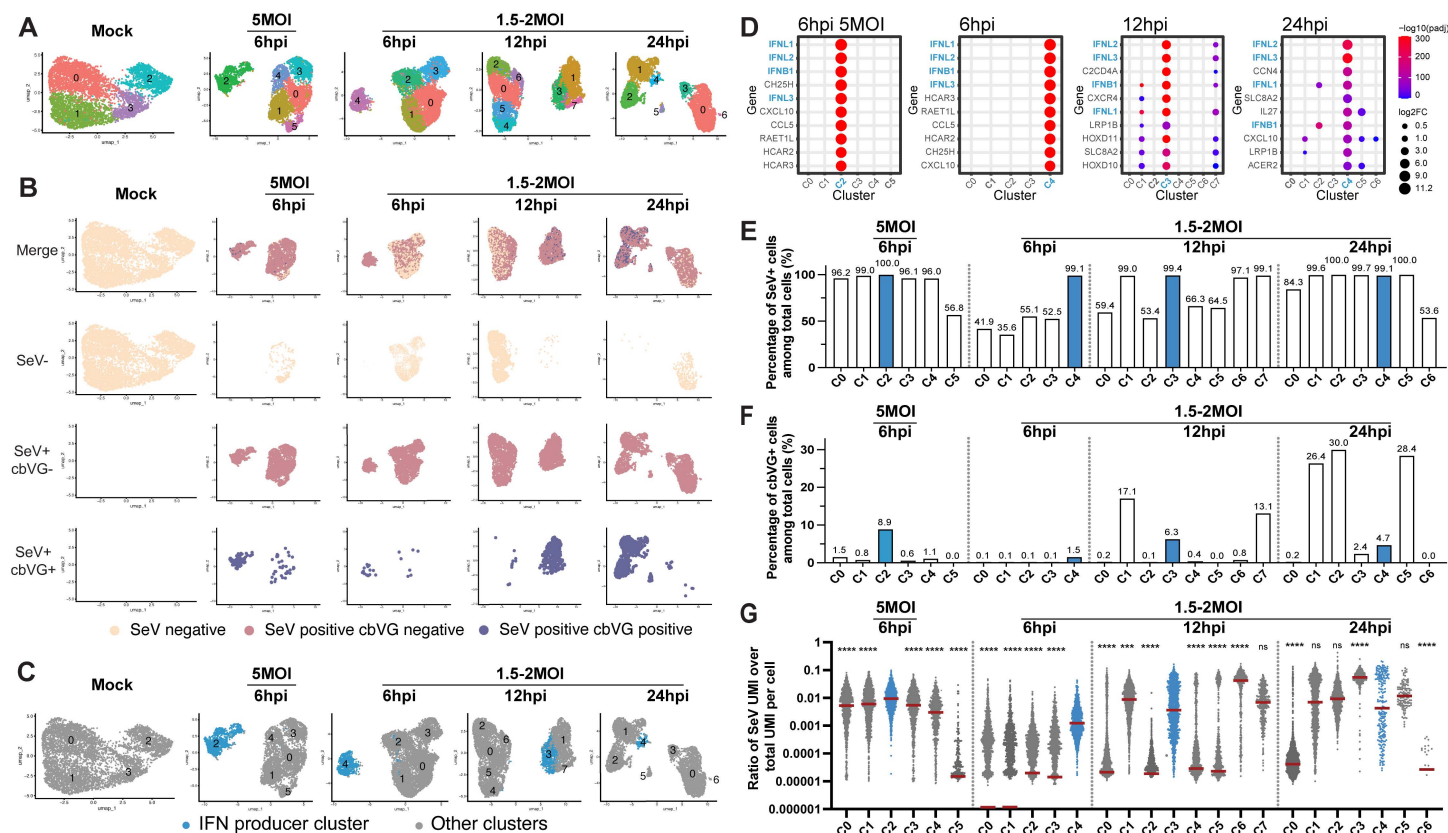


Fig 2. Unbiased clustering of each sample and identification of IFN-producing clusters in SeV-infected cells. (A) UMAPs showing unbiased clustering of mock and SeV-infected A549 cells at different infection conditions (5 MOI or 1.5-2 MOI) and time points (6, 12, and 24 hpi). Each color represents a distinct cell cluster in each sample. Clusters in different samples are not comparable based on their color. (B) UMAPs of SeV⁻, SeV⁺cbVG⁻, and SeV⁺cbVG⁺ cells in each sample. Cells are colored according to viral RNA detection status: SeV⁻ (beige), SeV⁺cbVG⁻ (pink), and SeV⁺cbVG⁺ (slate blue). (C) UMAPs showing IFN-producing (blue) clusters across mock and SeV-infected samples under different infection conditions (5 MOI or 1.5-2 MOI) and time points (6, 12, and 24 hpi). (D) Expression levels of the top 10 expressed marker genes of the IFN-producing clusters in different infection conditions. (E-F) Percentages of SeV⁺ cells (E) and cbVG⁺ cells (F) among total cells in each cluster. (G) Abundance of SeV reads in each cluster represented as the ratio of SeV UMI counts over total UMI counts per cell. Each dot represents a single cell. Statistical differences across all clusters in each sample were analyzed using the Kruskal-Wallis test. Following a significant overall test, pairwise comparisons were performed using Dunn's post-hoc test, in which each cluster was compared specifically against the IFN-producing cluster. Significance levels are indicated as ns for not significant and asterisks for significant differences (*** $p < 0.001$ and **** $p < 0.0001$).

<https://doi.org/10.1371/journal.ppat.1014184.g002>

distinct IFN-producing cluster was identified under all conditions, characterized by expression of IFNB1 and IFNL1–3 among the top marker genes (Fig 2C–2D). Although this cluster contained a high proportion of SeV⁺ cells, other clusters with similarly high infection rate did not express IFNs, indicating that infection alone is insufficient to initiate interferon production (Fig 2D–2E) and confirming published reports of a the lack of correlation between viral load and IFN expression [6,7,21,53].

SeV cbVG⁺ cells were consistently present within the IFN-producing cluster across MOIs and time points. At 6 hpi, the IFN-expressing cluster showed a higher average viral load than other clusters and higher counts of cbVGs (Fig 2D–2G) consistent with the requirement of viral replication machinery for cbVG accumulation [3,24,54]. This association diminished at later timepoints, likely explained by cbVG-driven interference with viral replication [3,55]. By 12 hpi, some cbVG⁺ clusters lacked IFN expression (Fig 2E–2G) likely reflecting host-mediated shutdown of interferon signaling and/or the transition of infected cells into different transcriptional states. Notably, many SeV⁺ cells across clusters had similar or higher total viral RNA abundance than cells in the IFN cluster, indicating that cbVG detection does not simply reflect viral load and demonstrating that cbVGs, rather than viral genome abundance, are the strongest correlate with interferon expression during early infection. Together, these results support data from bulk cell populations showing that cbVGs are the primary triggers of early antiviral programs [4,7–9,13,47,56,57] and establish that cbVGs are the initial drivers of antiviral signaling rather than it being a correlate of viral genome abundance.

Distinct and dynamic cellular transcriptional states are observed during SeV infection

To resolve cellular transcriptional profiles reflecting distinct cellular states across infection conditions, we performed integrated analysis followed by unsupervised clustering of SeV⁺ cells. This analysis revealed five distinct transcriptional profiles (C0–C4) that captured the dominant cellular programs present during the first 24 h of infection (Fig 3A). Each state was characterized by distinct gene expression programs (Fig 3B–3D), representing metabolic adaptation, cell-cycle regulation, epithelial secretory responses, interferon production, and inflammatory signaling. C0 expressed genes associated with metabolic and oxidative stress-response programs (e.g., HMOX1, SLC40A1, ACOX2), consistent with iron homeostasis, lipid metabolism, and redox balance (S3A Fig). C1 was enriched for cell-cycle and chromatin organization programs (e.g., HIST1H1B, CENPA, CCNA2), reflecting DNA replication and mitotic progression (S3B Fig). C2 was enriched in epithelial adhesion and secretory machinery genes (e.g., CEACAM6, MUC5AC, CLDN2, LCN2) together with cytokine-mediated signaling and immune activation programs (S3C Fig). C3 represented the interferon-producing state, defined by high expression of IFNB1 and IFNL1–3, and enriched for antiviral defense programs, including type I/III interferon signaling (S3D Fig). C4 exhibited inflammatory signaling modules, including chemokine induction and was enriched in NF-κB-associated signaling (e.g., CCL17, ICAM1, ACKR1) (S3E Fig). We then analyzed the expression levels of cluster-defining marker genes relative to mock and assessed the proportion of cells expressing them in each cluster at each experimental condition (Fig 3C–3D). All clusters contained more than 3,000 cells, enabling reliable assessment of gene expression levels and the proportion of marker-positive cells across conditions (S4A Fig). We observed that expression of metabolic and oxidative stress-related genes such as HMOX1, ACOX2, and CYP24A1, as well as cell-cycle and chromatin organization genes including HIST1H1B, CENPA, and CCNA2, were significantly downregulated in the IFN-producing (C3) and inflammatory (C4) clusters. In contrast, IFNL1, IFNL2, IFNL3, and IFNB1 were strongly upregulated in both C3 and C4).

We next quantified the distribution of cells across states over time (Fig 3E). While the proportion of C0 remained relatively constant, the other four transcriptional profiles displayed marked temporal dynamics. C1 was most abundant at 6 hpi and decreased thereafter, indicating suppression of proliferative programs as infection progressed. C2 diminished over time indicating reduction of cells with active adhesion and secretory pathways. C3 (interferon-producing) peaked at 12 hpi, consistent with early antiviral activation. By contrast, C4 expanded dramatically from ~1% at 6 hpi to 42% at 24 hpi, indicating a population shift toward late inflammatory programs. Analysis of SeV⁻ cells did not reveal comparable dynamics, as the top 20 marker genes in each cluster were predominantly interferon-stimulated genes (ISGs) and were upregulated

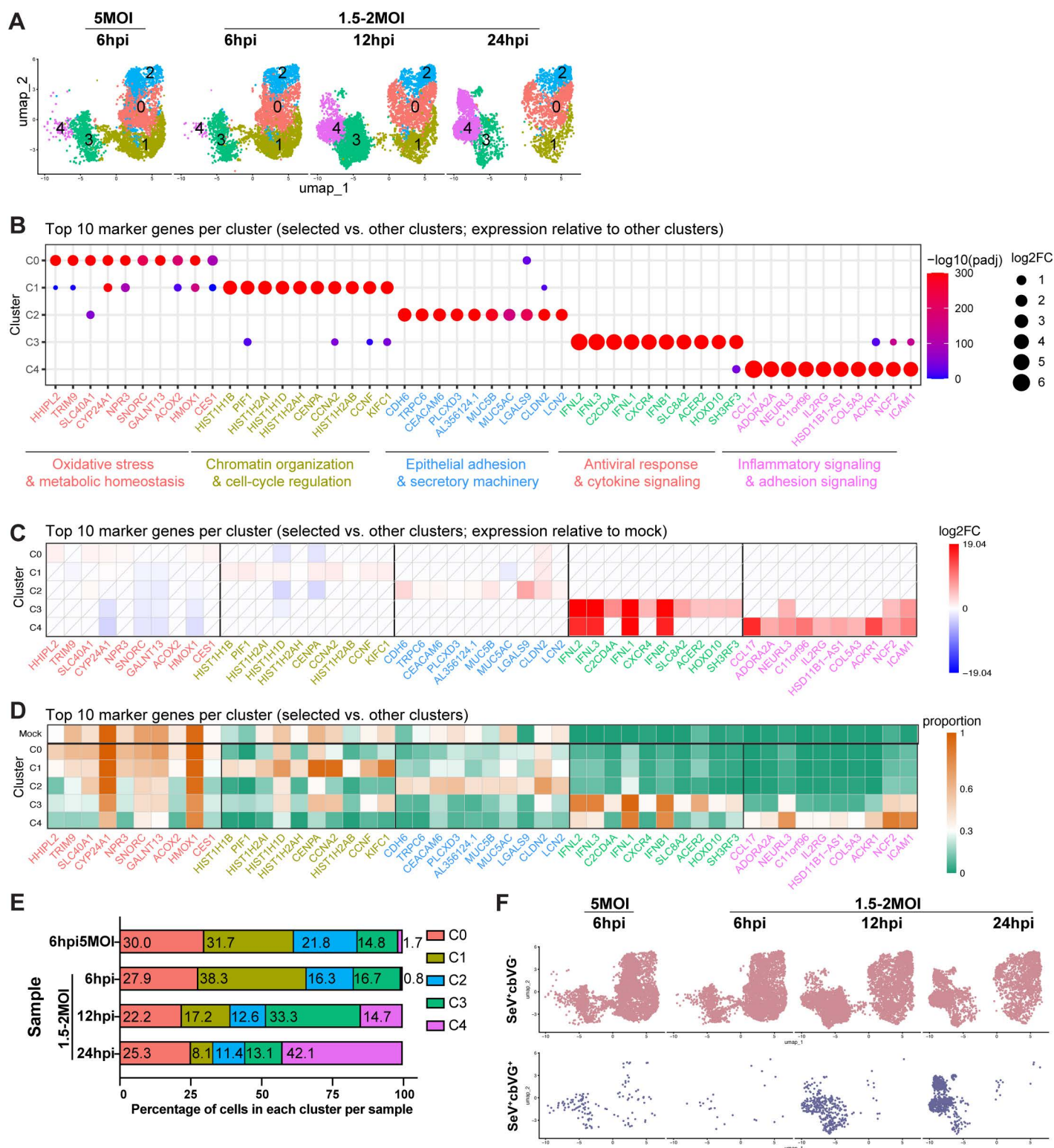


Fig 3. SeV⁺ cells can be divided into five populations during the first 24 h of infection based on transcriptome signatures. (A) Integrated analysis of all SeV⁺ cells from each sample. **(B)** Bubble plot showing the top 10 marker genes for each SeV⁺ cluster (C0-C4) identified relative to other clusters. Dot size represents log₂ fold change, and color indicates statistical significance (-log₁₀ adjusted p value). Functional categories of marker

genes are labeled below. **(C)** Heatmap showing the expression of the same top 10 marker genes in each cluster relative to mock-infected cells, with the color scale indicating log2 fold change. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. **(D)** Heatmap showing the proportion of cells expressing each marker gene across clusters and mock (uninfected sample), with the color scale representing expression proportion. **(E)** Percentage of cells in each population across samples. **(F)** UMAPs showing the distribution of SeV⁺ cells under different infection conditions (5 MOI or 1.5–2 MOI) and time points (6, 12, and 24 hpi). Separate panels show SeV⁺cbVG⁻ (pink) and SeV⁺cbVG⁺ (slate blue) populations individually.

<https://doi.org/10.1371/journal.ppat.1014184.g003>

across clusters relative to mock ([S5 Fig](#)), demonstrating that these transitions are specific to infected cells. To relate these transcriptional profiles to the presence of cbVGs, we mapped cbVG⁺ cells across clusters ([Figs 3E](#) and [S4B](#)). cbVG⁺ cells were concentrated in C3 and C4, with C3 containing most cbVG⁺ cells early in infection and C4 becoming the dominant cbVG⁺ population at 24 hpi. This pattern aligns with the temporal progression from early interferon induction (C3) to a later broader inflammatory response (C4).

Together, these analyses identify five functional transcriptional programs within a clonal population of infected cells, four of which change quantitatively within the first 24 h of infection. cbVG⁺ cells show primarily antiviral and inflammatory states, and their redistribution over time highlights a trajectory from early interferon production to late inflammatory signaling. Notably, metabolic and cell-cycle programs (C0 and C1), as well as cell adhesion and secretory programs (C2) were suppressed in C3 and C4 when comparing with mock, indicating that activation of antiviral programs coincides with downregulation of metabolic, proliferative and other processes, consistent with a cell-state transition toward a stress-adapted antiviral phenotype.

cbVGs impact the dynamics of the antiviral response within a population of infected cells

To more directly evaluate how cbVGs shape host responses within infected populations, we stratified SeV⁺ cells from each sample into cbVG⁺ and cbVG⁻ subgroups. The number of cells in each group and the SeV UMI counts per cells are shown in [S6 Fig](#). Then we compared transcriptional changes in cbVG⁺ and cbVG⁻ cells relative to mock controls ([S1–S2 Tables](#), and [S7 Fig](#)). cbVG⁺ cells displayed a greater number of differentially expressed genes than cbVG⁻ cells throughout the infection time course, except at the earliest stage of infection at low MOI ([Fig 4A–4B](#)), indicating an overall stronger transcriptional response associated with cbVG presence. When combining the top upregulated genes from all groups ([Fig 4C–4D](#)), interferons and chemokines (e.g., IFNB1, IFNL1–3, CCL5, CXCL10) were exclusively found in cbVG⁺ cells during early infection and became broadly expressed in both cbVG⁺ and cbVG⁻ cells at later time points. This observation is consistent with IFN priming and sensitization of bystander cells to respond to the infection. In contrast, a subset of genes, including CCL3 and GNG8, were induced only in cbVG⁺ cells at later timepoints, suggesting a temporal split on cbVG-driven transcriptional responses. Classic interferon-stimulated genes (ISGs), such as OASL, IFIT1, and ISG15, were induced in both cbVG⁺ and cbVG⁻ cells from 6 hpi onward, reflecting autocrine and paracrine responses to IFNs ([Fig 4C–4D](#)). A larger proportion of downregulated genes was observed at 12 and 24 hpi relative to 6 hpi ([Fig 4E–4F](#)), consistent with broad suppression of metabolic and proliferative programs as infection progressed.

To validate whether these genes were induced by cbVGs and not by the SeV infection alone, we quantified SeV NP, IFNL1, and IFIT1 mRNA expression in cells infected with either a cbVG-clean stock prepared by rescuing the virus by reverse genetics, a recently described cbVG-high stock (rS2) with a unique composition of cbVGs [14], or SeV Cantell that contains cbVG-546 ([Fig 4G–4I](#)). The dominant cbVG species present in these stocks are shown in [S8 Fig](#). NP levels remained high in cells infected with the clean stock across time points and MOIs, consistent with productive viral replication in the absence of pre-existing cbVGs ([Fig 4G](#)). IFNL1 and IFIT1 were not upregulated in infections with the cbVG-clean stock but showed robust induction in cells infected with cbVG-high stocks at multiple time points ([Fig 4H–4I](#)), confirming that their activation is driven by cbVG presence rather than the viral infection itself. These results provide orthogonal validation of the single-cell findings, confirming that key antiviral genes are selectively induced in

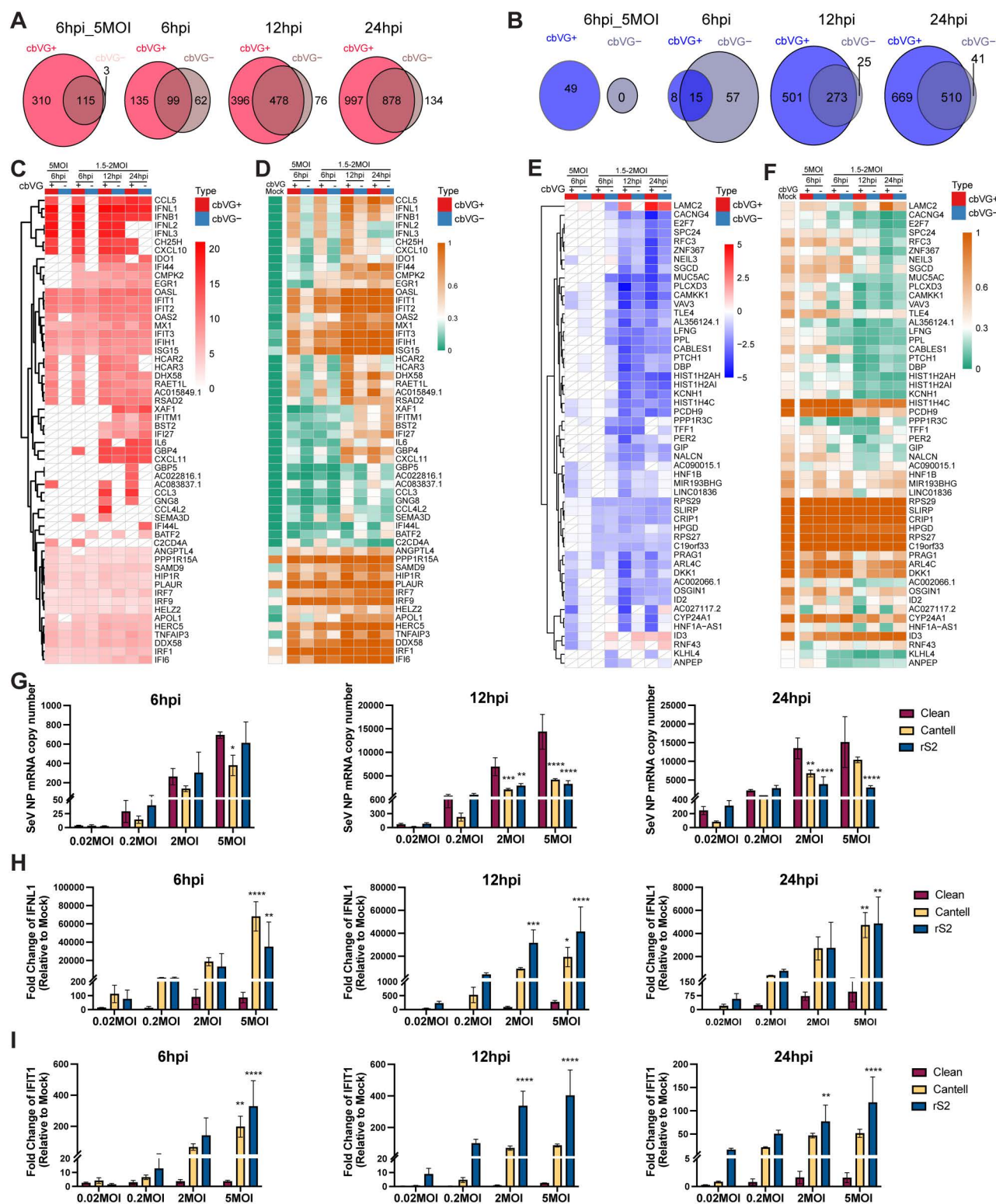


Fig 4. Viral RNA levels and cbVG-associated gene expression across infection conditions. (A-B) Venn diagrams showing the overlap of DEGs between cbVG⁺ and cbVG⁻ cells compared with mock-infected cells at 6hpi (5MOI), 6hpi, 12hpi, and 24hpi. Numbers indicate unique and shared genes between the two groups. (A) Upregulated genes (pct.cbVG⁺ > 0.3, padj < 0.01, and log2FC > 1). (B) Downregulated genes (pct.mock > 0.3, padj < 0.01, and

log₂FC < -1). **(C-F)** Heatmaps showing the expression levels (C and E) and the proportion of expressing cells (D and F) for the top 20 genes significantly upregulated (C-D) or downregulated (E-F) in cbVG⁺ or cbVG⁻ cell populations in each sample. The color scale indicates the average log₂ fold change (C and E) or the proportion of cells expressing each gene (**D and F**). White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. **(G)** SeV NP, **(H)** IFNL1, and **(I)** IFIT1 expression in A549 cells infected with SeV cbVG clean stock or cbVG-high stocks (Cantell and rS2) at indicated MOIs (0.02, 0.2, 2, and 5) and time points (6, 12, and 24 hpi). Expression of mRNA was calculated relative to the housekeeping index with GAPDH and β -actin. IFNL1 and IFIT1 expression represents fold change relative to mock-infected controls. Data represent means $\pm$ SD of three biological replicates. Data were analyzed using two-way ANOVA, followed by Šidák's multiple-comparisons test to compare Clean versus Cantell or rS2. Adjusted P values are shown. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.

<https://doi.org/10.1371/journal.ppat.1014184.g004>

the presence of cbVGs. Together with the transcriptomic data, they show that cbVGs determine both the timing and amplitude of early antiviral induction by triggering interferon and chemokine programs in a rare cbVG⁺ subpopulation, which then propagates to cbVG⁻ cells through paracrine signaling. Beyond this shared antiviral phase, cbVG⁺ cells sustain a distinct subset of late transcriptional programs, demonstrating that cbVGs also drive divergent cell-state outcomes later in infection.

cbVG induce distinct early and late transcriptional programs in infected cells

To more comprehensively define the transcriptional programs driven by cbVG, we then identified all genes that were significantly upregulated over mock in cbVG⁺ cells, but not cbVG⁻ cells, at different times post infection. We identified 43 genes upregulated at 6 hpi regardless of MOI (**Fig 5**). The genes in this category included interferons (IFNB1, IFNL1–3) and the chemokines CCL5 and CXCL10, which were strongly induced exclusively in cbVG⁺ cells (**Fig 5A**). In contrast, in cbVG⁻ cells these genes were not upregulated, despite the cells being infected. However, by 12 hpi, most of these genes were expressed in both cbVG⁺ and cbVG⁻ cells, and remained elevated at 24 hpi, indicating that the broader infected population express these antiviral effectors at later time points (**Fig 5A**). This temporal shift from cbVG-restricted to population-wide expression is consistent with enhanced responsiveness of cells to the infection, wherein IFNs and chemokines produced by a small cbVG⁺ subset signal surrounding cells to be primed to respond to the infection. We classified these early induced genes into seven functional programs based on known or predicted functions: interferon responses; chemokine signaling; metabolic and lipid signaling; complement and innate immune amplification; cell–cell interaction; transcriptional and stress-response regulation; and uncharacterized genes with predicted immune roles. These categories were curated by the authors based on initial gene ontology analysis. Of note, because we only considered genes expressed at both MOIs tested for a more robust analysis, we left out several genes that are detectable only at high MOI, including OAS2, TRAF1, and PAK3 (**S9 Fig**). Overall, these data show that early cbVG⁺ cells induce multifunctional antiviral responses rich in immune communication signals.

To validate these data, we next performed infections with a cbVG-clean stock, and two cbVG-high stocks, Cantell and rS2, and tested for gene expression by qPCR. Our data confirmed that early induced genes (**Fig 5B–5G**, **S10A–S10C Fig**) were detectable only in cbVG-high infections and were absent in infections with the cbVG-clean stock. These data mirror the scRNA-seq results and demonstrate that early antiviral gene induction is cbVG-dependent, is not specific to the cbVG-546 species, and can be induced by diverse cbVG populations. Together, these findings show that early antiviral signaling originates from a rare population of cbVG⁺ cells, which acts as the initial source of interferon and chemokine production and initiate a population-wide antiviral response through paracrine amplification.

To characterize the later cbVG-driven transcriptional programs, we analyzed genes that were upregulated only in cbVG⁺ cells at 12 and 24 hpi, after excluding those already induced at 6 hpi (**Fig 5H**). This analysis revealed 105 genes that define a late cbVG-dependent program. Among the top 41 genes (log₂FC > 2) were ATP4A, SCN3A, KRT17, and CCL3, which were highly and specifically enriched in cbVG⁺ cells at late time points. These late-phase cbVG⁺ programs encompass cell-surface receptor signaling, Ca²⁺-linked communication, cytoskeletal remodeling, transcriptional

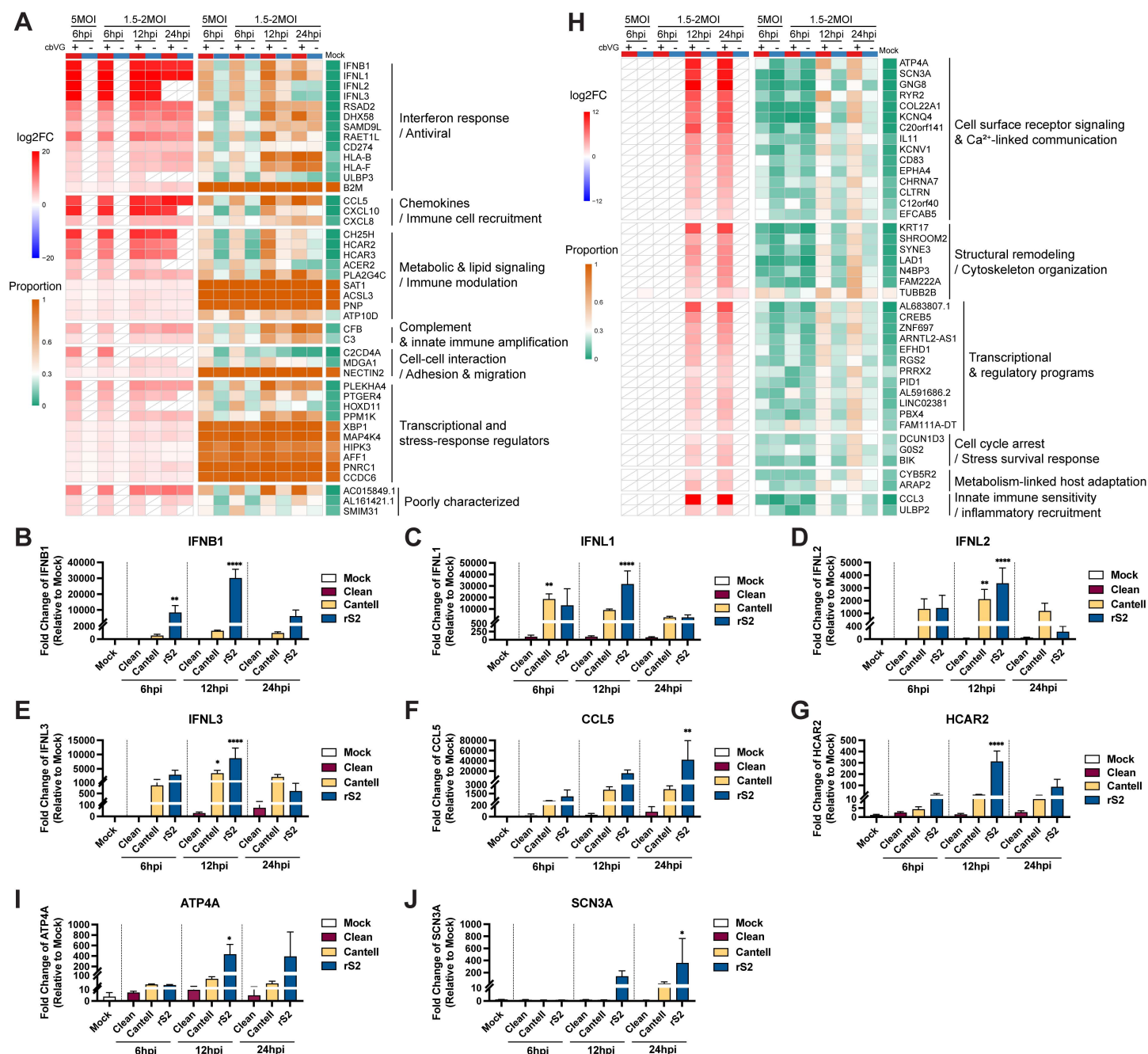


Fig 5. Functional classification and validation of genes upregulated exclusively in cbVG⁺ cells. (A) Heatmap showing genes specifically upregulated in cbVG⁺ but not cbVG⁻ cells at both 6hpi samples. Genes were grouped into seven functional categories based on their known or predicted biological roles. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing each gene. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. (B-G) Representative genes validated by qPCR. A549 cells were infected with SeV Cantell stocks containing different cbVG compositions: cbVG clean stock, Cantell: high level of cbVG-546, rS2: high levels of cbVGs other than cbVG-546. Each gene mRNA levels were quantified by qPCR and are presented as fold change relative to mock-infected samples. Cells collected at 6, 12, and 24 hpi. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test comparing each infection group with the mock control (n=3). Adjusted P values are shown. Significance P values are indicated as follows: P<0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****). (H) Heatmap showing genes specifically upregulated in cbVG⁺ but not cbVG⁻ cells at both 12hpi and 24hpi samples. Genes were grouped into six functional categories based on their known or predicted biological functions. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing each gene. White tiles

with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. (I–J) Representative genes validated by qPCR. A549 cells were infected with SeV Cantell stocks containing different cbVG compositions: cbVG clean stock, Cantell, and rS2. mRNA levels were quantified by qPCR and are presented as fold change relative to mock-infected samples. Cells were collected at 6, 12, and 24 hpi. Data represent mean $\pm$ SD from three biological replicates. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test comparing each infection group with the mock control (n=3). Adjusted P values are shown. Significance P values are indicated as follows: $P < 0.05$ (*), < 0.01 (**), < 0.001 (***), < 0.0001 (****).

<https://doi.org/10.1371/journal.ppat.1014184.g005>

regulation, stress-survival mechanisms, and innate immune sensitivity, indicating that cbVG⁺ cells undergo distinct cell-state transitions that are not shared by the broader population in the time frame analyzed. Genes that were transiently upregulated at 12 hpi or 24 hpi are shown in [S11 Fig](#). qPCR validation confirmed that these late-induced genes were not upregulated at 6 hpi but became strongly induced in infections with SeV Cantell and rS2 at later time points. These transcripts remained absent in infections with cbVG-clean stock ([Fig 5I–5J](#), [S10D–S10F Fig](#)). These results indicate that cbVG⁺ cells do not only initiate antiviral programs but later acquire a second wave of transcriptional responses that reflect cell-state remodeling beyond interferon defenses.

Several antiviral programs are shared across the infected cell population

We next defined the shared antiviral response by identifying genes that were upregulated in both cbVG⁺ and cbVG[−] cells throughout the infection course ([Fig 6](#)). We identified 70 genes that were significantly upregulated in both 6 hpi samples ([Fig 6A](#)), and nearly all remained elevated during 12 and 24 hpi, revealing a stable shared antiviral program. These broadly induced genes clustered into five expression programs based on temporal and functional features: (1) canonical ISGs; (2) interferon-responsive non-canonical genes; (3) NF- κ B-associated inflammatory and stress-responsive genes; (4) interferon signaling components; and (5) non-ISG transcriptional responses. At 6 hpi, cbVG⁺ cells expressed higher levels of these genes likely due to interferon signaling originating from the cbVG⁺ subset. By 12 and 24 hpi, these differences diminished as shared antiviral programs spread across the population. To assess whether broad induction of these genes is cbVG-dependent, we performed qPCR validation and found that OASL, IFIT1, HERC5, and TNFAIP3 were upregulated only in cbVG-high infections, but not in the cbVG-clean stock ([Fig 6B–6E](#)), demonstrating that even population-wide antiviral programs require cbVGs for initial induction. We also identified genes that were not broadly upregulated at 6 hpi but became strongly induced in both cbVG⁺ and cbVG[−] cells at 12 and 24 hpi ([Fig 6F](#)). The functions of these genes are mainly related to interferon response and innate antiviral defense, cell surface signaling and immune interaction, transcriptional and stress-response regulation, vesicle trafficking and autophagy, metabolic and mitochondrial adaptation, and other cellular responses, suggesting that these genes contribute to a broader and more coordinated phase of host adaptation at later stages of infection. qPCR validation shown genes CXCL11, IFI27, and ITGA2 were upregulated in all cbVG-high stock infected cells but not in cells infected with the cbVG-clean stock or in mock controls ([Fig 6G–6I](#)).

Together, these data define a shared antiviral program that becomes established across the infected population, but whose timing, magnitude, and breadth depend on early cbVG signals from a rare initiating subset of cells.

cbVG⁺ cells undergo late-stage suppression of structural and metabolic programs

We then examined downregulated genes to assess whether cbVGs repress cellular functions during infection. At 6 hpi, no genes were consistently downregulated across cbVG⁺ cells, however 44 genes were significantly downregulated specifically in cbVG⁺ cells at 12 and 24 hpi ([Fig 7A](#)). Some of these genes were also downregulated in the 6 hpi high-MOI condition, suggesting dose-dependent early repression. These repressed genes encompassed cell-surface signaling and adhesion, metabolic and mitochondrial adaptations, cytoskeletal organization, transcriptional and stress-response regulation, and vesicle trafficking, indicating that cbVG⁺ cells undergo structural and metabolic

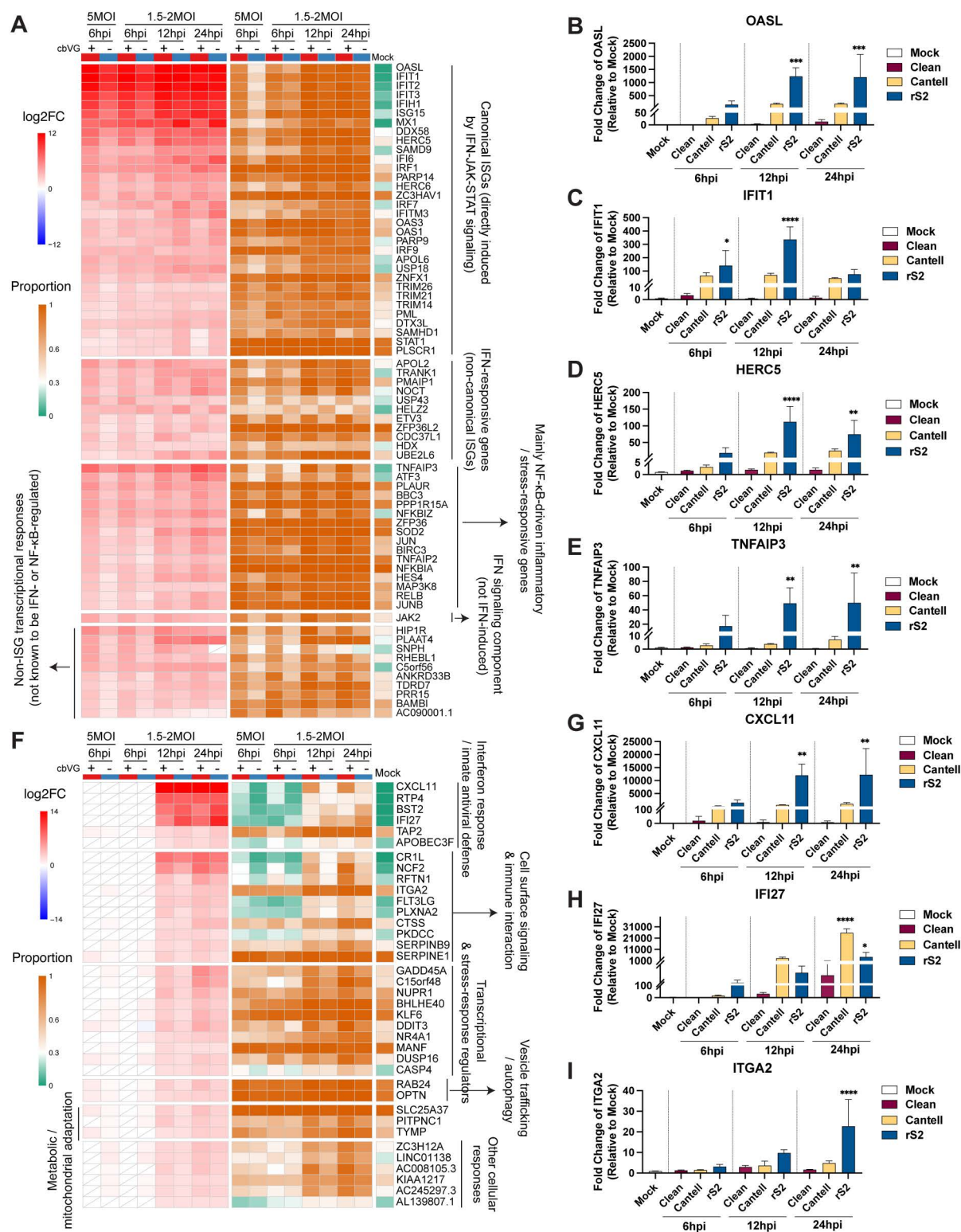


Fig 6. Functional classification and validation of genes upregulated in cbVG⁺ and cbVG⁻ cells. (A) Heatmap showing genes specifically upregulated in cbVG⁺ and cbVG⁻ cells at both 6hpi samples. Genes were grouped into five functional categories based on their transcriptional induction mechanisms. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing

each gene. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. **(B-E)** Representative genes validated by qPCR. A549 cells were infected with SeV Cantell or rS2. mRNA levels were quantified by qPCR and are presented as fold change relative to mock-infected samples. Cells were collected at 6, 12, and 24 hpi. Data represent mean $\pm$ SD from three biological replicates. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test comparing each infection group with the mock control ($n=3$). Adjusted P values are shown. Significance P values are indicated as follows: $P<0.05$ (*), <0.01 (**), <0.001 (***), <0.0001 (****). **(F)** Heatmap showing genes specifically upregulated in cbVG⁺ and cbVG⁻ cells at both 12 hpi and 24 hpi, but not at 6 hpi. Genes were grouped into seven functional categories based on their known or predicted biological functions. The red-blue color scale indicates the log2 fold change (log2FC), while the orange-green color scale represents the proportion of cells expressing each gene. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. **(G-I)** Representative genes validated by qPCR. A549 cells were infected with SeV Cantell or rS2. mRNA levels were quantified by qPCR and are presented as fold change relative to mock-infected samples. Cells were collected at 6, 12, and 24 hpi. Data represent mean $\pm$ SD from three biological replicates. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test comparing each infection group with the mock control ($n=3$). Adjusted P values are shown. Significance P values are indicated as follows: $P<0.05$ (*), <0.01 (**), <0.0001 (****).

<https://doi.org/10.1371/journal.ppat.1014184.g006>

remodeling at late stages of infection. This shift may reflect transition from active antiviral defense toward cellular recovery mechanisms. We attempted to validate downregulated genes by qPCR, however the low levels of expression of these genes and the impact of the observed upregulation of some of the genes in cbVG⁻ cells make this validation approach impossible.

Across all SeV-infected groups, we identified another set of genes with reduced expression at 12 and 24 hpi relative to mock (Fig 7B). Most of these genes showed stronger downregulation in cbVG⁺ cells, suggesting that cbVGs enhance late-stage suppression of cellular programs. These findings indicate that cbVG⁺ cells undergo late-phase structural and metabolic remodeling, reflecting a transition away from proliferative and epithelial programs during infection.

A multilayered model of cbVG-driven host responses

Collectively, our single-cell analyses reveal a tightly regulated, multilayered architecture of host responses driven by cbVGs (Fig 8): Layer 1 (Origin layer, early): a rare cbVG⁺ population initiates interferon and chemokine programs, acting as the source of antiviral signaling within the infected population. Layer 2 (Amplified layer): these antiviral signals propagate to cbVG⁻ cells and establish a population-wide antiviral state, dominated by canonical ISGs and innate signaling programs. Layer 3 (Divergent layer, late): cbVG⁺ cells subsequently acquire distinct late-phase transcriptional programs related to cell-state remodeling, and stress adaptation. Layer 4 (Repression layer): both cbVG⁺ and cbVG⁻ cells undergo late-stage downregulation of structural, metabolic, and developmental programs, with stronger repression in cbVG⁺ cells. Thus, cbVGs determine not only the magnitude of antiviral responses, but also the timing, distribution, and differentiation of cell states during infection. These findings provide a mechanistic framework for understanding how cbVGs shape infection outcomes by orchestrating sequential transcriptional programs and guiding cell-state transitions within infected populations.

Discussion

In this study, we developed a single-cell RNA sequencing workflow combined with a computational approach that enables direct detection of cbVGs and the host responses they induce at single-cell resolution. This strategy revealed that cbVGs orchestrate the antiviral response and drive cell-state transitions during infection, providing the first high resolution view of cbVG-associated cellular response during a viral infection. Rather than uniformly activating antiviral programs, cbVGs initiate interferon and chemokine signaling in a rare subset of infected cells, which subsequently transmit these signals to surrounding cells and establish a second wave of antiviral transcriptional responses. This layered architecture indicates that cbVGs determine the timing, distribution, and magnitude of antiviral programs across the infected population.

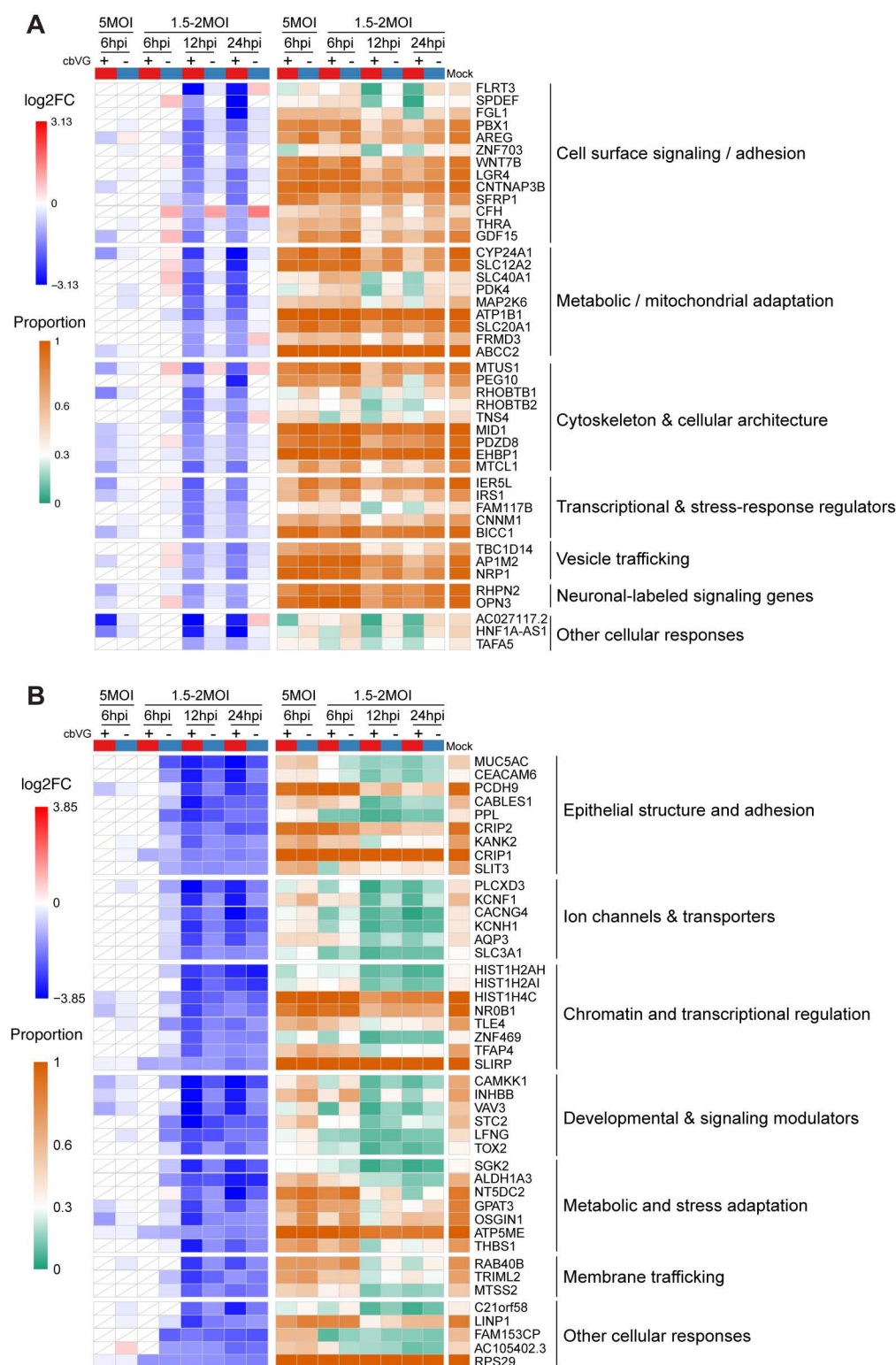


Fig 7. Functional classification and validation of genes downregulated at late infection. (A) Heatmap showing genes specifically downregulated in cbVG⁺ but not cbVG⁻ cells at both 12hpi and 24hpi samples. Genes were grouped into seven functional categories based on their known or predicted biological functions. The red-blue color scale indicates the log2 fold change (log2FC), while the orange-green color scale represents the proportion of

cells expressing each gene. **(B)** Heatmap showing genes specifically downregulated in cbVG⁺ and cbVG⁻ cells at both 12hpi and 24hpi samples but not common at 6hpi. Genes were grouped into seven functional categories based on their known or predicted biological functions. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing each gene. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds.

<https://doi.org/10.1371/journal.ppat.1014184.g007>

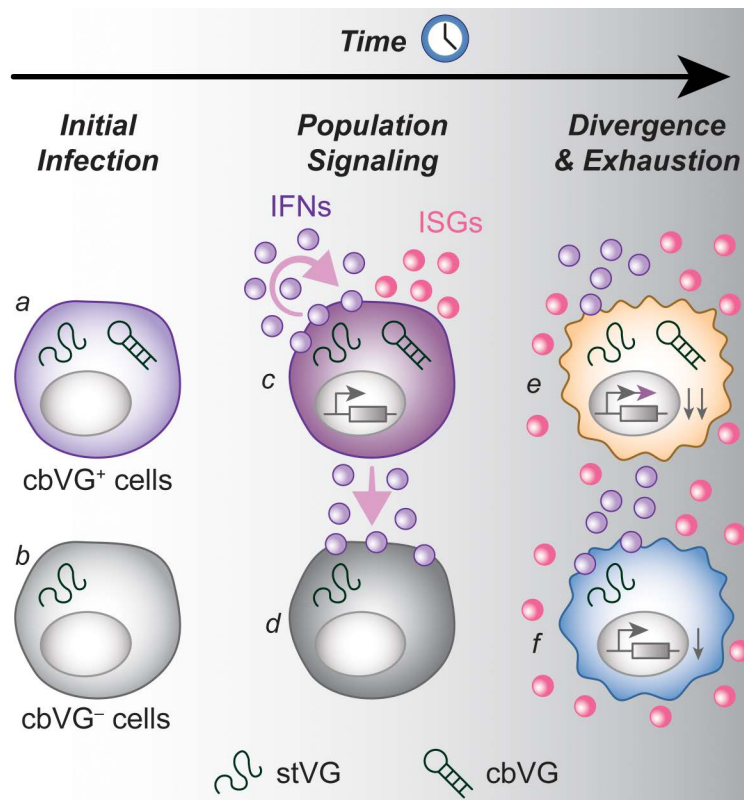


Fig 8. Model summarizing the role of cbVGs in driving antiviral responses during infection. At initial infection, there are two types of infected cells: cbVG⁺ cells (a) and cbVG⁻ cells (b). cbVG⁺ cells act as early initiators of innate immune signaling, inducing strong interferon (IFN) and interferon-stimulated gene (ISG) responses (c) and transmitting antiviral signals to neighboring cbVG⁻ cells through paracrine signaling (d). As infection progresses, IFN and ISG expression becomes broadly activated across the infected cell population (e, f). At later stages, cbVG⁺ cells additionally initiate unique transcriptional programs (e), while global transcriptional repression is observed in all infected cells and is stronger in cbVG⁺ cells (e) than in cbVG⁻ cells (f).

<https://doi.org/10.1371/journal.ppat.1014184.g008>

Previous bulk measurements have shown that cbVGs can dominate antiviral signaling [9,13,35,58], but these approaches mask cell-to-cell heterogeneity and obscure whether cbVGs simply correlate with antiviral activation or directly instruct population-level responses. At single-cell resolution, our analysis resolved how population-level antiviral responses are organized, revealing that cbVG⁺ cells act as early initiators, whereas cbVG⁻ cells function as receivers that later transition toward inflammatory programs. The identification of several classical interferon-stimulated genes in our IFN responder clusters indicates that IFN protein was produced in these infections. Integrated clustering further revealed five transcriptional states whose relative abundance shifts over time. Using differential expression analyses of cbVG⁺ and cbVG⁻ cells, we identified gene sets that reflect distinct layers of cbVG-associated responses. Together, these single-cell-resolved findings provide a framework for interpreting the widespread presence of cbVGs in clinical infection across diverse RNA viruses [2,7,16,18,19,21,57,59]. While the single-cell analysis tracks cbVG-546, the host transcriptional programs identified likely reflect responses to cbVGs more generally based on our qPCR validation using stocks

with distinct cbVG species. Because upon scRNA-seq cbVG abundance is quantified using UMIs rather than total amplified reads, cbVG reads at single-cell resolution are less than in bulk RNA sequencing and direct comparisons between data sets need to take this fact into consideration.

While our data support a dominant role for cbVG-associated signaling in initiating antiviral responses, we cannot fully exclude the possibility that other viral PAMPs, such as leader, trailer RNAs or readthrough RNAs, may also contribute to innate immune activation during infection. Nevertheless, IFN expression in our dataset is largely restricted to cbVG-high cells at early time points despite the expected presence of even higher levels of leader, trailer RNAs or readthrough RNAs in infected cells lacking cbVGs, suggesting that cbVGs are the primary triggers of the observed antiviral transcriptional programs. An important question that remains unresolved is how many cbVGs are required for triggering of the anti-viral response and whether there is a minimal ratio of cbVG/standard viral genomes required. Regrettably, our scRNA-seq method does not provide us with this quantitative power. In scRNA-Seq, viral cbVG transcripts are typically detected at much lower abundance compared to bulk RNA-seq due to the limited capture efficiency and sequencing depth per cell. For this reason, detection of at least one cbVG UMI was considered sufficient to classify a cell as cbVG-positive in the current study and it is unlikely that all cbVG molecules present in a cell are detected. Under these conditions, false negatives are more likely than false positives, preventing us from providing reasonable estimates on viral genomes numbers.

In the reported experiments we infected cells with a SeV Cantell stock with a high content of cbVGs. An important question that arises is whether detection of cbVGs incoming in the virus stock during infection is necessary for the triggering of the anti-viral response, and whether cbVGs generated *de novo* in the infected cell are equally immunostimulatory. To answer these questions, we need to refer to data collected through several years of work. Paramyxoviruses, including SeV maintain their genomes tightly packed in nucleoproteins and only genomes with a length of an exact multiple of six (known as the “rule of six”) replicate efficiently. This strict length is because each nucleoprotein binds 6 nucleotides leaving no exposed RNA. In published experiments using UV-inactivated viruses, we have demonstrated that UV inactivated virus does not induce IFN expression. Similarly, adding UV inactivated defective particles to a cbVG-free live virus stock does not lead to IFN induction [6,13,60]. These data suggest that the immunostimulatory motif of cbVGs [10,13] is only exposed during their replication, once nucleoproteins are removed by polymerase co-factors to give the polymerase access to the viral RNA. UV inactivation impedes unpacking of the RNA and its replication, keeping the immunostimulatory RNA hidden from viral sensors. These results do not imply that only pre-packaged cbVGs trigger IFN. Instead, they suggest that replication of cbVGs, whether newly generated, or pre-existing in the stock, is required to expose their immunostimulatory domains to cellular sensors. However, *de novo* generated cbVGs take time and specific conditions to accumulate at significant amounts. In the conditions of the experiments presented in this paper (MOI = 1.5–5, 24hpi), the most abundant cbVGs will be cbVG-546, which is carried in the virus stock, as can be seen in the gel in [S8 Fig](#) (SeV Cantell). Therefore, following cbVG-546 represents the only cbVG present at significant levels during our experiments, but it does not exclude the possibility that *de novo* generated cbVGs are also immunostimulatory.

In vivo and clinical observations have established correlative relationships between cbVG abundance and disease outcome, viral loads, or antiviral responses [7,21,57,59], without resolving how cbVGs shape host responses at the level of individual cells. cbVGs have been shown to spread within infected tissues [18] and to associate with interferon and inflammatory signatures [7,21], but whether these effects reflect direct sensing of cbVGs or secondary propagation of antiviral signals has remained unclear. Here, to directly define cbVG-associated host responses, we intentionally adopted a simplified experimental system using the human lung epithelial cell line A549. While this approach enabled us to isolate transcriptional programs tightly linked to cbVGs, it does not capture the full complexity of clinical infection. *In vivo*, virus infection including cbVG activity is likely influenced by host-specific variables such as age, sex, genetic background, and prior immune history [61–65], as well as by the functional specialization of distinct cell types as they play different roles during virus infection [66,67].

These considerations raise important unresolved questions regarding cbVG biology *in vivo*, including whether cbVGs preferentially arise or accumulate within specific cell types, and whether distinct cell populations exhibit differential sensitivity or transcriptional responses to cbVG-mediated signaling. Addressing these questions will require extending our findings along two complementary directions. First, the cbVG-associated gene sets identified here provide a resource for analyzing clinical samples, enabling evaluation of how these transcriptional programs correlate with cbVG abundance and clinical outcomes across patient cohorts. Second, *in vivo* infection models using cbVG-high viral stocks will be essential to determine whether cbVGs are preferentially enriched within particular cell types and to directly compare cbVG⁺ and cbVG⁻ responses within defined cellular compartments. Notably, because sensitive and comprehensive detection of diverse cbVG species in complex tissues remains technically challenging, the gene signatures defined in this study may also serve as indirect readouts for cbVG activity. For example, cells co-expressing multiple genes that were detected only in cbVG⁺ cells could be classified as functionally cbVG-positive, enabling identification of putative false-negative cells in which cbVGs are present but fall below current detection thresholds. Together, these approaches will help bridge single-cell mechanistic insights with the complexity of cbVG biology *in vivo* and in clinical infections. While our study focuses on epithelial cells, the conceptual separation of cbVG-initiating and cbVG-receiving states provides a framework that can be tested across additional cell types *in vivo*, including immune and stromal populations.

Our study also highlights technical considerations that will be important for future work. Mock controls were collected at 6 hpi, and epithelial gene expression can vary over time due to basal cycling, stress responses, or culture conditions, which may influence identification of late-regulated genes. Also, to minimize the effects of extensive cell death at later stages of infection, our analyses were limited to samples collected up to 24 hpi; therefore, potential cellular states emerging at later time points, such as the establishment of persistent infection in cbVG⁺ cells [35], were not captured and should be considered in future studies. Additionally, standard scRNA-seq platforms capture only short sequences near the template-switch oligo, which limits detection of diverse cbVG species and can lead to false-negative classification of some cbVG⁺ cells. This is not an issue for SeV Cantell infection, which is dominated by a single abundant cbVG species (cbVG-546), but it becomes problematic for other virus stocks that produce a wider variety of cbVGs. Despite these limitations, the cbVG-associated gene signatures identified here provide a complementary framework for interpreting cbVG activity at the single-cell level, particularly in contexts where direct cbVG detection remains incomplete. As sequencing technologies continue to advance, long-read single-cell approaches may enable full-length capture of cbVGs and direct linkage of individual cbVG species to their induced transcriptional programs. Combining these technologies with cell-type-specific models will be essential for defining how cbVG generation, accumulation, and host responses differ across diverse tissues and for resolving the molecular basis of how cbVGs shape infection dynamics in complex biological settings.

Materials and methods

Cells and virus

A549 cells (ATCC, #CCL-185) were cultured in tissue culture medium (Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, #11965092) supplemented with 10% fetal bovine serum (FBS) (Sigma, #F0926), gentamicin 50ug/ml (ThermoFisher, #15750060), L-glutamine 2 mM (Invitrogen, #G7513) and sodium pyruvate 1 mM (Invitrogen, #25-000-C1) at 5% CO₂ 37°C. Cells were treated with mycoplasma removal agent (MP Biomedical, #3050044) and tested monthly for mycoplasma contamination using the MycoAlert Plus mycoplasma testing kit (Lonza, #LT07-318). SeV Cantell strain was propagated in 10-day-old embryonated chicken eggs (Charles River Laboratories) for 40 hours, as previously described [60]. SeV clean stocks and cbVG-high stock rS2 were rescued from cDNA using reverse genetics and amplified in LLCMK2 cells. Their cbVG abundance and species composition were characterized as previously described [14].

Virus infection and samples preparation for scRNA-Seq

Infections were performed after washing the cells once with PBS and incubating them with virus diluted in infection media (DMEM, 35% bovine serum albumin (Sigma, #A7979), penicillin-streptomycin (Gibco, #15140–122), and 5% NaHCO₃ (Gibco, #25080094)) at 37°C for 1 hour, with gentle shaking every 15 minutes. Cells were then washed twice with PBS and supplemented with fresh infection media. The infected cells were incubated at 37°C and harvested at 6, 12, and 24 hpi. Mock-infected cells (infection media only) were used as controls. For each infection, two replicates were included. Infected or mock-infected cells were collected by detachment with Trypsin-EDTA 0.05% (Thermo, # 25300054) and subjected to live/dead staining (Thermo Fisher, # 65-0866-14), followed by flow cytometry sorting to isolate live cells.

Library preparation and sequencing

Live cells were resuspended in 1% BSA in 1×PBS at a concentration of 1,000 cells/μL, and 15 μL from each sample was used for library preparation. For each sample, 35.2 μL of a virus-specific RT primer (10 uM; 5'-AAGCAGTGG TATCAACGCAGAGTACTGCTCCTCAGGGTGGATACT-3') was added in combination with the conventionally used oligo-dT primer at an equal volume and concentration. Thus, both the virus-specific RT primer and oligo-dT were included during reverse transcription to simultaneously capture viral genomes/cbVGs and poly-adenylated host and viral transcripts. Library preparation was performed following the 10x Genomics Chromium Next GEM Single Cell 5' v2 protocol. Samples were sequenced on an Illumina NovaSeq 6000 to generate 150 bp paired-end reads, targeting 700 million reads per sample, at the Genome Access Technology Center (GTAC) core sequencing facility at Washington University. Raw sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1390230.

Upstream computational analyses of scRNA-Seq data

Raw Illumina sequencing data were aligned to the human reference genome (GRCh38) using Cell Ranger (v7.2.0) to generate gene count matrices, which were used to create Seurat objects for host gene expression analysis. Reads that did not map to the human genome were extracted from the Cell Ranger output and aligned to the SeV genome using Cell Ranger with a custom-built reference. In parallel, cbVG junction reads were identified using VODKA2 following specific preprocessing steps [30]. Briefly, the template switching oligo (TSO) sequence was removed from R1 using fastx_trimmer, and Illumina adapters, oligo(dT) tails, and 5' overhang sequences were trimmed from both R1 and R2 using cutadapt. The resulting SeV standard genome and cbVG read counts per barcode were then added to the Seurat objects as metadata. Doublet cells were identified and removed using the scDblFinder R package, and low-quality cells were filtered based on quality control (QC) metrics using a sample specific approach. Filtering thresholds were defined as the median ± 3 × median absolute deviation (MAD) for key QC metrics, utilizing the outlier detection logic implemented in the isOutlier() function from the scater package [68]. Specifically, cells were excluded if their mitochondrial transcript proportion was upward outlier, or if their total RNA counts or number of detected features were downward outliers. All downstream analyses and visualization were performed in R (v4.4 or 4.5).

Clustering in seurat

Per-sample clustering: Each sample was processed independently to capture sample-specific heterogeneity. After normalization and variance stabilization using SCTransform(), principal component analysis (PCA) was conducted, followed by dimensionality reduction via Uniform Manifold Approximation and Projection (UMAP). Clustering was performed using the Seurat FindClusters() function at a resolution of 0.2 or 0.4. Differential gene expression (DGE) analysis was performed to characterize distinct cell populations within each sample. Integrated clustering: To identify shared transcriptional programs across timepoints, we performed integrated clustering on specific cell subsets (e.g., all SeV-positive cells or across all

SeV-negative cells) from infected samples. Datasets were integrated using the standard Seurat anchor-based strategy. Following SCTransform on individual samples, we identified common integration features across datasets, ran PrepSCTIntegration() and identified anchors using FindIntegrationAnchors(). After merging the data with IntegrateData(), the integrated object was then subjected to PCA and UMAP reduction. Shared cell states were identified using a clustering resolution of 0.2 and DGE analysis was conducted across these integrated clusters to identify conserved transcriptional modules.

Gene expression analysis of cbVG⁺ and cbVG⁻ cells relative to mock

To identify cbVG-specific transcriptional programs, SeV-infected cells were stratified into cbVG⁺ and cbVG⁻ populations based on the detection of cbVG junction sequences. DGE analysis was performed using Seurat FindMarkers() to compare cbVG⁺ versus mock and cbVG⁻ versus mock uninfected cells independently. Due to technical limitations in cbVG detection, a small number of cbVG⁺ cells may be included in the cbVG⁻ population; therefore, we applied a percentage threshold for DGE analysis. Genes were defined as significantly upregulated in cbVG⁺ or cbVG⁻ cells if they met the following criteria: detected in >30% of cells in the population (min.pct > 0.3) and adjusted p-value < 0.01, and log2 fold-change > 1. Conversely, genes were defined as significantly downregulated relative to mock if they were detected in > 30% of mock cells, had an adjusted p-value < 0.01, and log2 fold-change < -1. This analytical design distinguishes cbVG-dependent responses from changes driven by SeV infection alone.

Genes validation by RT-qPCR

A549 cells were infected with SeV stocks (clean stock, Cantell, and rS2) at a defined MOI and harvested at 6, 12, and 24 hpi. Each infection condition included three biological replicates. Total RNA from infected and control samples was extracted using a KingFisher instrument and the MagMAX mirVana Total RNA Isolation Kit (Thermo Fisher, #A27828) following the manufacturer's guidelines. A total of 150 ng RNA was used for cDNA synthesis with the High-Capacity RNA-to-cDNA Kit (Thermo Fisher, #18080051). qPCR was performed using SYBR Green (Thermo Fisher, #A46113) and 5 μM of forward and reverse primers for each gene on an Applied Biosystems QuantStudio 5 machine. Primers used for qPCR are listed in [S3 Table](#). Relative gene expression was normalized to human GAPDH and human beta-actin expression as described previously [69]. qPCR reactions were performed in technical triplicates and data represent the mean of biological replicates.

Supporting information

S1 Fig. Visualization of the generation and replication of standard viral genomes (stVG) and copyback viral genomes (cbVG). (A) stVG replication. The viral RNA-dependent RNA polymerase (RdRp, blue) synthesizes a full-length stVG using the negative-sense genomic RNA as a template. (B) cbVG generation. During replication, RdRp detaches from the template at a break point (blue) and rejoins the nascent strand at a rejoin position (green), forming a cbVG with complementary ends and a unique junction region (red). White stripes mark the template sequence used after the rejoin point, and green stripes indicate the complementary sequence at the 3' end of the newly formed cbVG. (C) cbVG replication. cbVGs are replicated by RdRp. Pink and blue stripes indicate sequences complementary to the 3' and 5' ends of the cbVG template, respectively. (TIF)

S2 Fig. Workflow for detecting and integrating cbVGs from single-cell RNA-seq data. Raw Illumina sequencing data were processed through parallel bioinformatic workflows for host and viral analyses. On the left, reads were aligned to the human reference genome using Cell Ranger, followed by generation of the host gene count matrix and doublet removal (scDbtFinder). Non-human reads were extracted and re-aligned to the SeV genome to obtain SeV reads and barcodes.

On the right, reads were pre-processed to remove host sequences, TSO, and Illumina adapters, then analyzed with VODKA2 for cbVG junction detection to obtain cbVG reads and barcodes. Host, SeV, and cbVG read count matrices were integrated by cell barcode using Seurat for multimodal analysis and visualization.

(TIF)

S3 Fig. Gene Ontology (GO) enrichment analysis of SeV⁺ cell clusters. Bubble plots showing the top enriched GO biological processes for marker genes of each SeV⁺ cluster (C0–C4). The x-axis indicates fold enrichment, the y-axis lists enriched GO terms, dot size represents the number of genes associated with each term, and color indicates the adjusted p-value ($-\log_{10} p$). (A) C0, (B) C1, (C) C2, (D) C3, and (E) C4.

(TIF)

S4 Fig. Number of cells in integrated analysis. (A) Number of SeV⁺ cells in each cluster. (B) Number of cbVG⁺ cells in each cluster.

(TIF)

S5 Fig. SeV⁻ cells were divided into five populations based on host gene transcription. (A) Integrated analysis of all SeV⁻ cells from each sample. (B) Percentage of cells in each population across samples. (C) Number of SeV⁻ in each cluster. (D) Total UMI counts per cell. The red line indicates the median. (E) Heatmap of the top 20 most upregulated genes from each cluster compared with mock. Average log₂ fold change is shown as the scale bar. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds.

(TIF)

S6 Fig. SeV UMI counts per cell in cbVG⁺ and cbVG⁻ groups. Scatter plots showing SeV UMI counts per cell in cbVG⁺ (red) and cbVG⁻ (blue) populations at different condition. Each dot represents a single cell, and horizontal bars indicate median values. Numbers above each group denote the number of cells analyzed.

(TIF)

S7 Fig. Differential gene expression in cbVG⁺ and cbVG⁻ cells compared with mock (uninfected) cells. Volcano plots showing differentially expressed genes between cbVG⁺ or cbVG⁻ cells and mock cells at the indicated infection stages: (A, B) 6 hpi 5MOI, (C, D) 6 hpi, (E, F) 12 hpi, and (G, H) 24 hpi. Each dot represents one gene, with the x-axis showing average log₂ fold change (avg_log2FC) and the y-axis showing $-\log_{10}$ adjusted p-value. Red dots indicate upregulated genes, and blue dots indicate downregulated genes in cbVG⁺ (left) or cbVG⁻ (right) SeV⁺ cells. For visualization, p values equal to 0 were replaced with one-tenth of the smallest non-zero p value across all genes.

(TIF)

S8 Fig. Dominant cbVG species present in the virus stocks used in this study. cbVG-specific and genomic PCR analysis of SeV Cantell and the high-MOI passaged stock rS2 from passages P1 (cbVG-clean stock) to P8. LLC-MK2 cells were infected at an MOI of 1.5 for 24 h prior to RNA extraction. SeV Cantell (C) contains the predominant cbVG-546 species. Mock (M) indicates uninfected cells. The asterisk marks rS2 P3, which corresponds to the rS2 stock used in this study.

(TIF)

S9 Fig. Expression levels and proportions of selected genes. Heatmap showing genes specifically upregulated ($\text{log}_2 > 2$) in cbVG⁺ 6hpi 5MOI samples. The red-blue color scale indicates the log₂ fold change (log2FC) (A), while the orange-green color scale represents the proportion of cells expressing each gene (B). White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds.

(TIF)

S10 Fig. Representative genes validated by qPCR. A549 cells were infected with SeV stocks containing different cbVG compositions: cbVG clean stock, Cantell, and rS2. Each gene mRNA levels were quantified by qPCR and are presented as fold change relative to mock-infected samples. Cells were collected at 6, 12, and 24 hpi. Data represent mean $\pm$ SD from three biological replicates. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test comparing each infection group with the mock control ($n=3$). Adjusted P values are shown. Significance P values are indicated as follows: $P<0.05$ (*), <0.01 (**), <0.001 (***), <0.0001 (****). (TIF)

S11 Fig. Expression levels and proportions of selected genes. (A-B) Heatmap showing genes specifically upregulated in cbVG⁺ 12hpi samples. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing each gene. (C-D) Heatmap showing genes specifically upregulated in cbVG⁺ 24hpi samples. The red-blue color scale indicates the log₂ fold change (log₂FC), while the orange-green color scale represents the proportion of cells expressing each gene. White tiles with diagonal lines indicate features that did not meet the predefined adjusted P-value and proportion thresholds. (TIF)

S1 Table. SeV⁺cbVG⁺ over Mock. DGE analysis using SeV⁺cbVG⁺ cells over mock uninfected cells. (XLSX)

S2 Table. SeV⁺cbVG⁻ over Mock. DGE analysis using SeV⁺cbVG⁻ cells over mock uninfected cells. (XLSX)

S3 Table. Primers. All primer sequences used in this paper. (XLSX)

Acknowledgments

We thank the Genome Technology Access Center (GTAC) and Jennifer Ponce at the McDonnell Genome Institute at Washington University School of Medicine in St. Louis for assistance with library preparation and sequencing. We thank Dr. Jan Carette and Dr. Allison Dupzyk from the Department of Microbiology and Immunology at Stanford University for their early technical advice on virus-specific primer design. We also thank the Siteman Cancer Center Flow Cytometry Core at Washington University School of Medicine in St. Louis for assistance with sorting infected cells.

Author contributions

Conceptualization: Yanling Yang, Carolina B. López.

Data curation: Yanling Yang, Emna Achouri.

Formal analysis: Yanling Yang, Emna Achouri.

Funding acquisition: Carolina B. López.

Investigation: Yanling Yang, Munyaradzi Tambo.

Methodology: Yanling Yang, Emna Achouri, Munyaradzi Tambo.

Project administration: Carolina B. López.

Resources: Carolina B. López.

Software: Yanling Yang, Emna Achouri.

Supervision: Carolina B. López.

Validation: Yanling Yang.

Visualization: Yanling Yang, Emna Achouri.

Writing – original draft: Yanling Yang.

Writing – review & editing: Yanling Yang, Emna Achouri, Munyaradzi Tambo, Carolina B. López.

References

- González Aparicio LJ, López CB, Felt SA. A virus is a community: diversity within negative-sense RNA virus populations. *Microbiol Mol Biol Rev*. 2022;86(3):e0008621. <https://doi.org/10.1128/mmb.00086-21> PMID: 35658541
- Li X, Ye Z, Plant EP. 5' copyback defective viral genomes are major component in clinical and non-clinical influenza samples. *Virus Res*. 2024;339:199274. <https://doi.org/10.1016/j.virusres.2023.199274> PMID: 37981214
- Vignuzzi M, López CB. Defective viral genomes are key drivers of the virus-host interaction. *Nat Microbiol*. 2019;4(7):1075–87. <https://doi.org/10.1038/s41564-019-0465-y> PMID: 31160826
- Johnston MD. The characteristics required for a Sendai virus preparation to induce high levels of interferon in human lymphoblastoid cells. *J Gen Virol*. 1981;56(Pt 1):175–84. <https://doi.org/10.1099/0022-1317-56-1-175> PMID: 6170730
- van den Hoogen BG, van Boheemen S, de Rijck J, van Nieuwkoop S, Smith DJ, Laksono B, et al. Excessive production and extreme editing of human metapneumovirus defective interfering RNA is associated with type I IFN induction. *J Gen Virol*. 2014;95(Pt 8):1625–33. <https://doi.org/10.1099/vir.0.066100-0> PMID: 24760760
- Tapia K, Kim W-K, Sun Y, Mercado-López X, Dunay E, Wise M, et al. Defective viral genomes arising in vivo provide critical danger signals for the triggering of lung antiviral immunity. *PLoS Pathog*. 2013;9(10):e1003703. <https://doi.org/10.1371/journal.ppat.1003703> PMID: 24204261
- Sun Y, Jain D, Koziol-White CJ, Genoyer E, Gilbert M, Tapia K, et al. Immunostimulatory defective viral genomes from respiratory syncytial virus promote a strong innate antiviral response during infection in mice and humans. *PLoS Pathog*. 2015;11(9):e1005122. <https://doi.org/10.1371/journal.ppat.1005122> PMID: 26336095
- Strahle L, Garcin D, Kolakofsky D. Sendai virus defective-interfering genomes and the activation of interferon-beta. *Virology*. 2006;351(1):101–11. <https://doi.org/10.1016/j.virol.2006.03.022> PMID: 16631220
- Yount JS, Gitlin L, Moran TM, López CB. MDA5 participates in the detection of paramyxovirus infection and is essential for the early activation of dendritic cells in response to Sendai Virus defective interfering particles. *J Immunol*. 2008;180(7):4910–8. <https://doi.org/10.4049/jimmunol.180.7.4910> PMID: 18354215
- Xu J, Mercado-López X, Grier JT, Kim W, Chun LF, Irvine EB, et al. Identification of a natural viral RNA motif that optimizes sensing of Viral RNA by RIG-I. *mBio*. 2015;6(5):e01265-15. <https://doi.org/10.1128/mBio.01265-15> PMID: 26443454
- Baum A, Sachidanandam R, García-Sastre A. Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. *Proc Natl Acad Sci U S A*. 2010;107(37):16303–8. <https://doi.org/10.1073/pnas.1005077107> PMID: 20805493
- Runge S, Sparrer KMJ, Lässig C, Hembach K, Baum A, García-Sastre A, et al. In vivo ligands of MDA5 and RIG-I in measles virus-infected cells. *PLoS Pathog*. 2014;10(4):e1004081. <https://doi.org/10.1371/journal.ppat.1004081> PMID: 24743923
- Mercado-López X, Cotter CR, Kim W-K, Sun Y, Muñoz L, Tapia K, et al. Highly immunostimulatory RNA derived from a Sendai virus defective viral genome. *Vaccine*. 2013;31(48):5713–21. <https://doi.org/10.1016/j.vaccine.2013.09.040> PMID: 24099876
- Yang Y, Wang Y, López CB. Variable and conserved features of copy-back viral genome populations generated de novo during Sendai virus infection. *J Virol*. 2026;100(2):e0180525. <https://doi.org/10.1128/jvi.01805-25> PMID: 41511081
- Pye SE, Achouri E, Yang Y, Musa A, López CB. Validation of diverse and previously untraceable Sendai virus copyback viral genomes by direct RNA sequencing. *J Virol*. 2025;99(8):e0089425. <https://doi.org/10.1128/jvi.00894-25> PMID: 40742273
- Cattaneo R, Rebmann G, Schmid A, Baczko K, ter Meulen V, Billeter MA. Altered transcription of a defective measles virus genome derived from a diseased human brain. *EMBO J*. 1987;6(3):681–8. <https://doi.org/10.1002/j.1460-2075.1987.tb04808.x> PMID: 3582370
- Pfaller CK, Mastorakos GM, Matchett WE, Ma X, Samuel CE, Cattaneo R. Measles Virus Defective Interfering RNAs Are Generated Frequently and Early in the Absence of C Protein and Can Be Destabilized by Adenosine Deaminase Acting on RNA-1-Like Hypermutations. *J Virol*. 2015;89(15):7735–47. <https://doi.org/10.1128/JVI.01017-15> PMID: 25972541
- Taye B, Yousaf I, Navaratnarajah CK, Schroeder DC, Pfaller CK, Cattaneo R. A measles virus collective infectious unit that caused lethal human brain disease includes many locally restricted and few widespread copy-back defective genomes. *J Virol*. 2024;98(11):e0123224. <https://doi.org/10.1128/jvi.01232-24> PMID: 39431848
- Sidhu MS, Crowley J, Lowenthal A, Karcher D, Menonna J, Cook S, et al. Defective measles virus in human subacute sclerosing panencephalitis brain. *Virology*. 1994;202(2):631–41. <https://doi.org/10.1006/viro.1994.1384> PMID: 8030228
- Felt SA, Achouri E, Faber SR, López CB. Accumulation of copy-back viral genomes during respiratory syncytial virus infection is preceded by diversification of the copy-back viral genome population followed by selection. *Virus Evol*. 2022;8(2):veac091. <https://doi.org/10.1093/ve/veac091> PMID: 36325033

21. Felt SA, Sun Y, Jozwik A, Paras A, Habibi MS, Nickle D, et al. Detection of respiratory syncytial virus defective genomes in nasal secretions is associated with distinct clinical outcomes. *Nat Microbiol*. 2021;6(5):672–81. <https://doi.org/10.1038/s41564-021-00882-3> PMID: [33795879](#)
22. Sun Y, Kim EJ, Felt SA, Taylor LJ, Agarwal D, Grant GR, et al. A specific sequence in the genome of respiratory syncytial virus regulates the generation of copy-back defective viral genomes. *PLoS Pathog*. 2019;15(4):e1007707. <https://doi.org/10.1371/journal.ppat.1007707> PMID: [30995283](#)
23. Linder A, Bothe V, Linder N, Schwarzlmüller P, Dahlström F, Bartenhagen C, et al. Defective interfering genomes and the full-length viral genome trigger RIG-I after infection with vesicular stomatitis virus in a replication dependent manner. *Front Immunol*. 2021;12:595390. <https://doi.org/10.3389/fimmu.2021.595390> PMID: [33995343](#)
24. Peluso RW, Moyer SA. Viral proteins required for the in vitro replication of vesicular stomatitis virus defective interfering particle genome RNA. *Virology*. 1988;162(2):369–76. [https://doi.org/10.1016/0042-6822\(88\)90477-1](https://doi.org/10.1016/0042-6822(88)90477-1) PMID: [2829424](#)
25. O'Hara PJ, Nichol ST, Horodyski FM, Holland JJ. Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell*. 1984;36(4):915–24. [https://doi.org/10.1016/0092-8674\(84\)90041-2](https://doi.org/10.1016/0092-8674(84)90041-2) PMID: [6323026](#)
26. Boussier J, Munier S, Achouri E, Meyer B, Crescenzo-Chaigne B, Behillil S, et al. RNA-seq accuracy and reproducibility for the mapping and quantification of influenza defective viral genomes. *RNA*. 2020;26(12):1905–18. <https://doi.org/10.1261/rna.077529.120> PMID: [32929001](#)
27. Taylor A, Rosa C, Archetti M. Defective but promising: evaluating the utility of currently available bioinformatic pipelines for detecting defective viral genomes in RNA-Seq data. *J Gen Virol*. 2025;106(11):002176. <https://doi.org/10.1099/jgv.0.002176> PMID: [41247266](#)
28. Olmo-Uceda MJ, Muñoz-Sánchez JC, Lasso-Giraldo W, Arnau V, Díaz-Villanueva W, Elena SF. DVGfinder: A Metasearch Tool for Identifying Defective Viral Genomes in RNA-Seq Data. *Viruses*. 2022;14(5):1114. <https://doi.org/10.3390/v14051114> PMID: [35632855](#)
29. Muñoz-Sánchez JC, Olmo-Uceda MJ, Oteo JA, Elena SF. Quantifying defective and wild-type viruses from high-throughput RNA sequencing. *Bioinformatics*. 2024;40(11):btac651.
30. Achouri E, Felt SA, Hackbart M, Rivera-Espinal NS, López CB. VODKA2: a fast and accurate method to detect non-standard viral genomes from large RNA-seq data sets. *RNA*. 2023;30(1):16–25. <https://doi.org/10.1261/rna.079747.123> PMID: [37891004](#)
31. Russell AB, Elshina E, Kowalsky JR, Te Velthuis AJW, Bloom JD. Single-cell virus sequencing of influenza infections that trigger innate immunity. *J Virol*. 2019;93(14):e00500–19. <https://doi.org/10.1128/JVI.00500-19> PMID: [31068418](#)
32. Wang C, Forst CV, Chou T-W, Geber A, Wang M, Hamou W, et al. Cell-to-cell variation in defective virus expression and effects on host responses during influenza virus infection. *mBio*. 2020;11(1):e02880–19. <https://doi.org/10.1128/mBio.02880-19> PMID: [31937643](#)
33. González Aparicio LJ, Yang Y, Hackbart M, López CB. Copy-back viral genomes induce a cellular stress response that interferes with viral protein expression without affecting antiviral immunity. *PLoS Biol*. 2023;21(11):e3002381. <https://doi.org/10.1371/journal.pbio.3002381> PMID: [37983241](#)
34. Genoyer E, Kulej K, Hung CT, Thibault PA, Azarm K, Takimoto T, et al. The viral polymerase complex mediates the interaction of viral ribonucleoprotein complexes with recycling endosomes during sendai virus assembly. *mBio*. 2020;11(4):e02028–20. <https://doi.org/10.1128/mBio.02028-20> PMID: [32843550](#)
35. Xu J, Sun Y, Li Y, Ruthel G, Weiss SR, Raj A, et al. Replication defective viral genomes exploit a cellular pro-survival mechanism to establish paramyxovirus persistence. *Nat Commun*. 2017;8(1):799. <https://doi.org/10.1038/s41467-017-00909-6> PMID: [28986577](#)
36. Cristinelli S, Ciuffi A. The use of single-cell RNA-Seq to understand virus-host interactions. *Curr Opin Virol*. 2018;29:39–50. <https://doi.org/10.1016/j.coviro.2018.03.001> PMID: [29558678](#)
37. Ratnasiri K, Wilk AJ, Lee MJ, Khatri P, Blish CA. Single-cell RNA-seq methods to interrogate virus-host interactions. In: *Seminars in immunopathology*. Springer; 2023.
38. Suomalainen M, Greber UF. Virus infection variability by single-cell profiling. *Viruses*. 2021;13(8):1568. <https://doi.org/10.3390/v13081568> PMID: [34452433](#)
39. Chang J-T, Liu L-B, Wang P-G, An J. Single-cell RNA sequencing to understand host-virus interactions. *Virol Sin*. 2024;39(1):1–8. <https://doi.org/10.1016/j.virs.2023.11.009> PMID: [38008383](#)
40. Swaminath S, Russell AB. The use of single-cell RNA-seq to study heterogeneity at varying levels of virus-host interactions. *PLoS Pathog*. 2024;20(1):e1011898. <https://doi.org/10.1371/journal.ppat.1011898> PMID: [38236826](#)
41. Sun J, Vera JC, Drnevich J, Lin YT, Ke R, Brooke CB. Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. *PLoS Pathog*. 2020;16(7):e1008671. <https://doi.org/10.1371/journal.ppat.1008671> PMID: [32614923](#)
42. Bost P, Giladi A, Liu Y, Bendjelal Y, Xu G, David E, et al. Host-viral infection maps reveal signatures of severe COVID-19 patients. *Cell*. 2020;181(7):1475–1488.e12. <https://doi.org/10.1016/j.cell.2020.05.006> PMID: [32479746](#)
43. Bruins LJM, Müller M, Schipper JG, Rabouw HH, Boersma S, van Kuppeveld FJM, et al. Antiviral responses are shaped by heterogeneity in viral replication dynamics. *Nat Microbiol*. 2023;8(11):2115–29. <https://doi.org/10.1038/s41564-023-01501-z> PMID: [37814072](#)
44. Cui C, Tang X, Xing J, Sheng X, Chi H, Zhan W. Single-cell RNA-seq revealed heterogeneous responses and functional differentiation of hemocytes against white spot syndrome virus infection in *Litopenaeus vannamei*. *J Virol*. 2024;98(3):e0180523. <https://doi.org/10.1128/jvi.01805-23> PMID: [38323810](#)
45. Killip MJ, Young DF, Gatherer D, Ross CS, Short JAL, Davison AJ, et al. Deep sequencing analysis of defective genomes of parainfluenza virus 5 and their role in interferon induction. *J Virol*. 2013;87(9):4798–807. <https://doi.org/10.1128/JVI.03383-12> PMID: [23449801](#)

46. Calain P, Monroe MC, Nichol ST. Ebola virus defective interfering particles and persistent infection. *Virology*. 1999;262(1):114–28. <https://doi.org/10.1006/viro.1999.9915> PMID: [10489346](#)
47. Welch SR, Tilston NL, Lo MK, Whitmer SLM, Harmon JR, Scholte FEM, et al. Inhibition of nipah virus by defective interfering particles. *J Infect Dis*. 2020;221(Suppl 4):S460–70. <https://doi.org/10.1093/infdis/jiz564> PMID: [32108876](#)
48. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: a brief overview. *Clin Transl Med*. 2022;12(3):e694. <https://doi.org/10.1002/ctm2.694> PMID: [35352511](#)
49. Kumari P, Kaur M, Dindhoria K, Ashford B, Amarasinghe SL, Thind AS. Advances in long-read single-cell transcriptomics. *Hum Genet*. 2024;143(9–10):1005–20. <https://doi.org/10.1007/s00439-024-02678-x> PMID: [38787419](#)
50. Dupzyk AJ, Waldman BS, Zengel J, Carette JE. Dissecting the host determinants of flavivirus infection using QIC-seq. *bioRxiv*. 2025.
51. Zhou Y, Liu L, Yang W, Wu Y, Zhong C, Liu Y, et al. A novel ZIKV-targeted scRNA-seq method for precise quantification of ZIKV RNA. *J Virol*. 2025;99(10):e0111425. <https://doi.org/10.1128/jvi.01114-25> PMID: [40990513](#)
52. Zanini F, Robinson ML, Croote D, Sahoo MK, Sanz AM, Ortiz-Lasso E, et al. Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue. *Proc Natl Acad Sci U S A*. 2018;115(52):E12363–9. <https://doi.org/10.1073/pnas.1813819115> PMID: [30530648](#)
53. Killip MJ, Young DF, Ross CS, Chen S, Goodbourn S, Randall RE. Failure to activate the IFN- β promoter by a paramyxovirus lacking an interferon antagonist. *Virology*. 2011;415(1):39–46. <https://doi.org/10.1016/j.virol.2011.03.027> PMID: [21511322](#)
54. Horikami SM, Curran J, Kolakofsky D, Moyer SA. Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. *J Virol*. 1992;66(8):4901–8. <https://doi.org/10.1128/JVI.66.8.4901-4908.1992> PMID: [1321276](#)
55. Huang AS, Baltimore D. Defective viral particles and viral disease processes. *Nature*. 1970;226(5243):325–7. <https://doi.org/10.1038/226325a0> PMID: [5439728](#)
56. Ho T-H, Kew C, Lui P-Y, Chan C-P, Satoh T, Akira S, et al. PACT- and RIG-I-dependent activation of type I interferon production by a defective interfering RNA derived from measles virus vaccine. *J Virol*. 2015;90(3):1557–68. <https://doi.org/10.1128/JVI.02161-15> PMID: [26608320](#)
57. Dong X, Garcia-Dorival I, Rickett NY, Bosworth A, Smither S, Dowall S, et al. Discontinuous EBOV RNA synthesis events in patients with Ebola virus disease and their relationship to viral load and outcome of infection. *J Virol*. 2025;99(12):e0082625. <https://doi.org/10.1128/jvi.00826-25> PMID: [41217172](#)
58. Genoyer E, López CB. Defective viral genomes alter how sendai virus interacts with cellular trafficking machinery, leading to heterogeneity in the production of viral particles among infected cells. *J Virol*. 2019;93(4):e01579-18. <https://doi.org/10.1128/JVI.01579-18> PMID: [30463965](#)
59. Addetia A, Phung Q, Bradley BT, Lin MJ, Zhu H, Xie H, et al. In vivo generation of BK and JC polyomavirus defective viral genomes in human urine samples associated with higher viral loads. *J Virol*. 2021;95(12):e00250-21. <https://doi.org/10.1128/JVI.00250-21> PMID: [33827948](#)
60. Yount JS, Kraus TA, Horvath CM, Moran TM, López CB. A novel role for viral-defective interfering particles in enhancing dendritic cell maturation. *J Immunol*. 2006;177(7):4503–13. <https://doi.org/10.4049/jimmunol.177.7.4503> PMID: [16982887](#)
61. Jacobsen H, Klein SL. Sex differences in immunity to viral infections. *Front Immunol*. 2021;12:720952. <https://doi.org/10.3389/fimmu.2021.720952> PMID: [34531867](#)
62. Scepanovic P, Alanio C, Hammer C, Hodel F, Bergstedt J, Patin E, et al. Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. *Genome Med*. 2018;10(1):59. <https://doi.org/10.1186/s13073-018-0568-8> PMID: [30053915](#)
63. Ohuma EO, Okiro EA, Ochola R, Sande CJ, Cane PA, Medley GF, et al. The natural history of respiratory syncytial virus in a birth cohort: the influence of age and previous infection on reinfection and disease. *Am J Epidemiol*. 2012;176(9):794–802. <https://doi.org/10.1093/aje/kws257> PMID: [23059788](#)
64. Piasecka B, Duffy D, Urrutia A, Quach H, Patin E, Posseme C, et al. Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proc Natl Acad Sci U S A*. 2018;115(3):E488–97. <https://doi.org/10.1073/pnas.1714765115> PMID: [29282317](#)
65. Kachuri L, Francis SS, Morrison ML, Wendt GA, Bossé Y, Cavazos TB, et al. The landscape of host genetic factors involved in immune response to common viral infections. *Genome Med*. 2020;12(1):93. <https://doi.org/10.1186/s13073-020-00790-x> PMID: [33109261](#)
66. Revie D, Salahuddin SZ. Human cell types important for hepatitis C virus replication in vivo and in vitro: old assertions and current evidence. *Virology*. 2011;8:346. <https://doi.org/10.1186/1743-422X-8-346> PMID: [21745397](#)
67. Fay EJ, Aron SL, Macchietto MG, Markman MW, Esser-Nobis K, Gale M Jr, et al. Cell type- and replication stage-specific influenza virus responses in vivo. *PLoS Pathog*. 2020;16(8):e1008760. <https://doi.org/10.1371/journal.ppat.1008760> PMID: [32790753](#)
68. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*. 2017;33(8):1179–86. <https://doi.org/10.1093/bioinformatics/btw777> PMID: [28088763](#)
69. Garcia GL, Valenzuela A, Manzoni T, Vaughan AE, López CB. Distinct chronic post-viral lung diseases upon infection with influenza or parainfluenza viruses differentially impact superinfection outcome. *Am J Pathol*. 2020;190(3):543–53. <https://doi.org/10.1016/j.ajpath.2019.11.003> PMID: [31866346](#)