Loss of Sendai virus C protein leads to accumulation of RIG-I immunostimulatory defective interfering RNA

Maria Teresa Sánchez-Aparicio,1,2 Dominique Garcin,3 Charles M. Rice,4 Daniel Kolakofsky,3 Adolfo García-Sastre1,2,5,* and Alina Baum1,*†

Abstract
Retinoic acid inducible gene (RIG-I)-mediated innate immunity plays a pivotal role in defence against virus infections. Previously we have shown that Sendai virus (SeV) defective interfering (DI) RNA functions as an exclusive and potent RIG-I ligand in DI-RNA-rich SeV-Cantell infected cells. To further understand how RIG-I is activated during SeV infection, we used a different interferon (IFN)-inducing SeV strain, recombinant SeVΔC, which, in contrast to SeV-Cantell is believed to stimulate IFN production due to the lack of the SeV IFN antagonist protein C. Surprisingly, we found that in SeVΔC-infected cells, DI RNAs also functioned as an exclusive RIG-I ligand. Infections with wild-type SeV failed to generate any RIG-I-associated immunostimulatory RNA and this correlated with the lack of DI genomes in infected cells, as well as with the absence of cellular innate immune responses. Supplementation of the C protein in the context of SeVΔC infection led to a reduction in the number of DI RNAs, further supporting the potential role of the C protein as a negative regulator of DI generation and/or accumulation. Our findings indicate that limiting DI genome production is an important function of viral IFN antagonist proteins.

INTRODUCTION
Effective innate immune responses are critical for the control and clearance of pathogenic infections. Antiviral innate immunity is composed of specialized immune cells, such as macrophages and NK cells, as well as cell-intrinsic antiviral mechanisms that can be found in all cell types. Cell-intrinsic responses depend on the recognition of viral pathogen-associated molecular patterns (PAMPs) within target cells and the initiation of signalling cascades, leading to the establishment of an antiviral state and the production of many pro-inflammatory cytokines, including type I interferon (IFN). Viral recognition within target cells is accomplished by pattern-recognition receptors (PRRs), proteins that recognize conserved structures within an invading pathogen, allowing for self versus non-self discrimination [1]. Two major families of PRRs are responsible for the detection of RNA viruses and activation of type I IFN: toll-like receptors (TLRs) and RIG-I-like receptors (RLRs). While antiviral TLRs are commonly found in the endosomal compartments of immune cells, RLRs are ubiquitously expressed in the cytoplasm of the majority of cells. As such, RLR receptors are positioned to activate the very first wave of immune responses to an invading viral pathogen. The RLR family is composed of three PRRs: RIG-I, MDA5 and LGP2. Both RIG-I and MDA5 initiate innate immune signalling upon recognition of viral RNA in the cytoplasm and generate the primary inflammatory response to viral infections [2]. RIG-I is activated by binding to its viral RNA ligand – an RNA molecule containing a 5′-tri/diphosphate with an adjacent blunt-ended, double-stranded RNA region [3–10].

During Sendai virus (SeV) infection, RIG-I becomes activated through binding to the SeV copy-back defective interfering (cbDI) RNA [11]. Generated by the viral polymerase during replication, DI RNAs are composed of a subgenomic RNA with intact promoters, and although they are unable to replicate on their own, DI RNAs interfere with viral replication by sequestering the viral replication machinery [12]. Virus stocks rich in DI RNAs are well known to induce
a high IFN response during infection [13, 14]. The cbDI structure, with a 5′-ppp and a perfectly complementary dsRNA region directly adjacent to it, possesses all the structural elements required for robust RIG-I activation. In previous studies we used a strain of SeV, SeV-Cantell, that is rich in DI particles, to identify RIG-I ligands, as this virus has been extensively characterized as a RIG-I-dependent potent IFN inducer [15, 16]. Identification of SeV DI RNA as an exclusive RIG-I ligand in a DI-rich infection naturally led to the question of what RNA molecule(s) may activate RIG-I during the course of infection with a DI-free virus. Wild-type viruses with functional IFN antagonists stay hidden from the immune system for an extensive period of time [17]. Viruses that are known to strongly and quickly activate IFN pathways tend to fall into two distinct categories, those with high DI composition and those lacking an IFN antagonist [18]. Since our previous work was performed with a virus containing high amounts of DI, we wanted to extend our studies to a virus deficient in its ability to suppress IFN activation. SeV lacking the C protein (SeVΔC) has previously been shown to strongly activate the IFN response through the RIG-I pathway [19]. Accordingly, we utilized this virus to identify RNA molecules which act as RIG-I ligands during infection with a virus defective in IFN antagonism.

The C ORF of SeV is encoded within the P gene, which contains overlapping reading frames for viral V and C proteins. The V protein of SeV is generated through mRNA editing and inhibits IFN activation by binding to MDA5 [20]. The SeV C protein is synthesized from the P mRNA through utilization of alternative ribosomal initiation sites. A nested set of four proteins are expressed from the C ORF: C, C, Y1, and Y2. The C proteins have been shown to have multiple roles in the viral lifecycle, including regulation of viral replication and transcription [21–26], suppression of apoptosis [27], enhancement of viral budding and assembly [28–30], and inhibition of IFN production and signalling [19, 31–33]. The role of the C proteins in inhibition of IFN responses has been attributed to their ability to inactivate STAT1 [32] as well as RIG-I [19, 33].

To our surprise, we found that despite growth conditions which normally inhibit DI genome accumulation, such as plaque purification and passaging at low multiplicity of infection, SeVΔC nevertheless contained a high abundance of cbDI RNAs and these molecules were again found to function as exclusive RIG-I ligands during infection [34]. Trans-supplementation of SeV C protein during infection led to significant reduction in DI genome levels, supporting the hypothesis that the C protein has a role in suppressing the generation and/or accumulation of cbDI RNAs.

RESULTS

RIG-I binds SeVΔC cbDI RNA

To isolate RIG-I RNA ligands generated during SeVΔC infection, A549 cells were infected at a multiplicity of infection (m.o.i) of 10 and the infection was allowed to proceed for 24 h. At this point the cells were lysed and endogenous RIG-I/RNA complexes were isolated as previously described [11]. The immunoprecipitated RNA was tested for its immunostimulatory activity through transfection into a 293T-IFN-sensitive response element (ISRE)-Fluc reporter cell line, expressing an ISRE promoter-driven firefly luciferase (Fluc) construct (Fig. 1a). Transfection of RIG-I-associated RNA resulted in robust activation of the Fluc reporter compared to the control immunoprecipitated (IP) RNA. To assess the biochemical nature of the RIG-I ligand, we subjected the isolated nucleic acid to RNase A digestion and calf alkaline phosphatase (CIP) treatment (Fig. 1b). Both treatments resulted in complete loss of immunostimulatory activity, confirming the nature of the RIG-I ligand as an RNA molecule with a functionally important phosphate group.

To identify the precise nature of the RNA species interacting with RIG-I we performed deep sequencing analysis on total IP RNA from both RIG-I and control IPs. The RNA was
sequenced on the Illumina HiSeq platform and sequencing reads were aligned to the Sendai-Z virus genome (Fig. 2). Analysis of RNAs specifically enriched in RIG-I IPs versus control IPs revealed two distinct regions mapping to the 5' end of the SeV genome (Fig. 2, bottom two panels). The high abundance of sequencing reads mapping to the 5' end of the viral genome was characteristic of high levels of copy-back DI RNA (cbDI), which we previously identified as a RIG-I ligand [11]. As with our previous report with SeV-Cantell, we did not find any RIG-I-specific enrichment within the rest of the SeV genome (Fig. 2, panel 2). To confirm the nature of RIG-I-associated RNA as cbDIs, we performed deep sequencing analysis on viral RNA isolated from purified SeVΔC stock as well as virus stocks of several

---

**Fig. 2.** SeVΔC RNAs enriched in RIG-I IPs. Total RNA isolated with RIG-I (red) or control (blue) IP was subjected to deep sequencing. RNA-Seq reads are mapped to their corresponding position within the SeV genome. The top panel illustrates the entire SeV genome. The bottom three panels zoom in on highlighted regions of the SeV genome. The number of RNA-Seq reads is given on the Y-axis. RIG-I-specific enrichment, corresponding to DI RNAs, can be seen in the bottom two panels. The bottom panel encompasses reads mapping to both DI1 and DI2. The second from the bottom panel corresponds to reads mapping to DI2.
other Sendai viruses (Table 1). Since SeV DIs are known to be packaged into virus particles, we wanted to confirm that the RNA species that we identified as RIG-I ligands behaved as classical DIs and not another type of aberrant viral replication product. In order to identify sequencing reads belonging to cbDI molecules we developed a computational approach based on the identification of reads spanning a fusion junction within the cbDI. Our method relies on identification of DI-specific sequencing reads which map to the unique point in the cbDI where the virus polymerase makes a junction between the 3’ and 5’ ends of the RNA. This junction corresponds to the portion of the cbDI at the meeting point between the single-stranded and double-stranded stretch and is not found in the full-length genome or L gene mRNA, allowing the identification of DI-specific sequencing reads. By utilizing an algorithm which identifies these types of junctions in RNA-Seq data, we can identify the abundance, exact position and size of cbDI molecules without any a priori knowledge about their sequence. Furthermore, this computational approach allows us to identify the presence of extremely rare DI molecules as long as even a single sequencing read spanning their junction point has been generated. In order to test the robustness and accuracy of our approach we performed sequencing of viral RNA from SeV-Cantell, a strain that contains a well-characterized cbDI species. Indeed, our computational method was able to identify the exact cbDI molecule known to be associated with SeV-Cantell. Table 1 shows the exact position of the fusion junction, as well as the cbDI relative abundance, and the length of the associated cbDI molecules, together with the length of the corresponding dsRNA stretch. For all identified DIs we confirmed that the resulting RNA molecule conforms to the rule of six, a hallmark of SeV genome replication [35]. When we performed the same analysis on the SeVAC stock we identified two molecules that were identical to those identified from our RIG-I pulldown data. These corresponded to 14258/15280 (DI2) and 14984/15304 (DI1) fusion points. DI1 was approximately 15 times more abundant than DI2, closely mirroring the ratio of these two RNAs in infected cells, and indicating packaging at similar efficiency (Fig. 3). Comparison of the relative RIG-I enrichment of DI1 and DI2 revealed that DI1 is enriched more than 30-fold above background, compared with a 7-fold enrichment for DI2 (Fig. 4). It is not clear whether the greater enrichment for DI1 is due to preferential binding of this RNA to RIG-I, or is simply based on increased abundance of this RNA in infected cells. Although DI1 is 15-fold more abundant in infected cells than DI2, we only saw a fourfold increase in its RIG-I binding compared to DI2. It is possible that DI2 may present a better ligand for RIG-I, based on the longer stretch of the dsRNA region compared to DI1 (104 nt compared to 80 nt). However, many other aspects of virus replication could play a role in RIG-I binding, such as the physical availability of the RNA for interaction with RIG-I. Further analysis of cbDI sequences found in SeV stocks revealed the presence of several other cbDI species in both SeV-Cantell and SeVAC. Interestingly, the most abundant SeV-Cantell cbDI (14931/15291) was also found in stocks of SeVAC, although at minimal amounts (Table 1). Vice versa, we were able to identify the prominent SeVAC DI (DI1) in stocks of SeV-Cantell, but again at very low levels. To eliminate the possibility that the low level of DI sequences common to all viral strains was the result of cross-contamination between the samples, we checked the sequence of the 14984/15302 DI fusion reads from SeV-Cantell and SeV-Z. Strain-specific sequence variation at positions 15 032–15 034 allowed us to differentiate between fusion reads originating from Cantell versus Z strains. We

**Table 1.** Description of cbDIs associated with various virus stocks and identified through RNA-Seq analysis

RNA-Seq fusion reads were used to deduce the exact nature of the cbDIs associated with each listed virus. The exact fusion point of the cbDI, its overall length, and the length of the dsRNA panhandle are listed. The percentage of reads supporting fusion is reflective of the overall prevalence of a particular cbDI within the total viral RNA pool.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Fusion point</th>
<th>DI length (nt)</th>
<th>dsRNA length (nt)</th>
<th>No. of fusion reads</th>
<th>No. of total viral reads</th>
<th>Percentage of reads supporting fusion point</th>
</tr>
</thead>
<tbody>
<tr>
<td>SeV-Cantell</td>
<td>14931/15291</td>
<td>546</td>
<td>92</td>
<td>121 248</td>
<td>23 510 595</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>14984/15304</td>
<td>480</td>
<td>80</td>
<td>148</td>
<td>336</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>14950/15278</td>
<td>540</td>
<td>106</td>
<td>49</td>
<td>869</td>
<td>0.005</td>
</tr>
<tr>
<td>SeV-ΔC</td>
<td>14258/15280 (DI1)</td>
<td>1230</td>
<td>104</td>
<td>8435</td>
<td>16 361 704</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>14931/15291</td>
<td>546</td>
<td>92</td>
<td>49</td>
<td>869</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>14984/15304 (DI2)</td>
<td>480</td>
<td>80</td>
<td>127 405</td>
<td>16 361 704</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>15178/15284</td>
<td>306</td>
<td>100</td>
<td>16 040 655</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>SeV-GFP-C</td>
<td>14931/15291</td>
<td>546</td>
<td>92</td>
<td>38</td>
<td>16 040 655</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>14984/15304</td>
<td>480</td>
<td>80</td>
<td>59</td>
<td>16 040 655</td>
<td>0.0004</td>
</tr>
<tr>
<td>SeV-GFP</td>
<td>14931/15291</td>
<td>546</td>
<td>92</td>
<td>24</td>
<td>5 770 180</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>14984/15304</td>
<td>480</td>
<td>80</td>
<td>19</td>
<td>5 770 180</td>
<td>0.0003</td>
</tr>
<tr>
<td>SeV-GFP-P6</td>
<td>14931/15291</td>
<td>546</td>
<td>92</td>
<td>49</td>
<td>17 478 494</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>14984/15304</td>
<td>480</td>
<td>80</td>
<td>33</td>
<td>17 478 494</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
could not perform the same analysis on the predominant Cantell DI species (14 931/15 291) because the genomic sequence encompassed by the fusion reads from that DI were identical between the Cantell and Z strains. The identification of identical cbDIs within two different strains of SeV indicates that cbDIs are prone to arise de novo at specific positions within the virus genome, likely through the influence of the underlying genomic sequence, which is identical at the fusion point between these two strains. The subsequent expansion of select cbDI species likely establishes a prominent DI associated with a particular SeV strain. In addition to these two species we found several other cbDIs present in low amounts within both SeV-Cantell and SeV-Z stocks. Out of all the identified cbDIs, the 14258/15280 (DI2) associated with SeV-D represented the longest species at 1230 nt. We performed a Western blot to detect the protein C expression in cells infected with the different viruses analysed in the experiments (Fig. 4b).

**Fig. 4.** (a) Quantitative enrichment of RIG-I binding by SeV RNA species. Ratios of RIG-I-associated reads over control IP reads were calculated. (b) 293 T cells were infected with the various Sendai viruses and total cell lysates were analysed by Western blot with SeV C protein antiserum 24 h post-infection. Bands corresponding to C protein, as well as P and V proteins, can be seen.

**DI-free SeV fails to activate RIG-I**

Our identification of DI RNA as a ligand for RIG-I in SeVΔC-infected cells led us to postulate that other recombinant viruses lacking functional IFN antagonists may induce the IFN response through DI accumulation. Since we wanted to examine which, if any, RNA molecules would interact with RIG-I during infection with a virus free of DI RNA, we decided to utilize a virus with wild-type-like
growth kinetics and IFN suppression which should be free of DI molecules. To this end we chose to perform experiments with SeV-GFP, a recombinant virus originating from the same Z strain as SeVAC and shown to behave in a way identical to a parental wild-type virus [36]. In order to confirm the DI-free nature of SeV-GFP we utilized the same computational approach as previously described. Analysis of viral RNA from SeV-GFP stock revealed a very low presence of both DI species associated with SeV-Cantell and SeVAC (Fig. 4a). Interestingly, when we performed the same analysis on SeV-GFP, which was passaged under conditions associated with DI accumulation, (i.e. six undiluted serial passages in embryonated eggs), we did not find any enrichment in cbDI levels, indicating that this particular virus is extremely resistant to DI accumulation. To find which RNA molecules may associate with RIG-I during SeV-GFP infection we performed RIG-I IPs in SeV-GFP-infected cells. Since levels of endogenous RIG-I in the absence of IFR3/7 activation or IFN signalling are very low, we decided to perform the immunoprecipitations in the context of RIG-I overexpression to make sure that any negative results would not be due to limited amounts of the endogenous RIG-I protein. Infections with SeV-GFP and SeVAC were performed in parallel with an m.o.i. of 1 and RIG-I/RNA complexes were isolated in the same manner as above, with the exception of using an anti-HA antibody instead of an anti-RIG-I antibody. The isolated RNA was again tested for immunostimulatory activity in 293T-ISRE-Fluc cells (Fig. 5a). Unlike infections with SeVAC, infections with SeV-GFP failed to produce any RIG-I-associated immunostimulatory RNA, suggesting that a wild-type SeV lacking DI RNA does not normally produce any RIG-I-specific PAMPs. Consistent with this observation, infection of the 293T ISRE-RFP cells with SeV-GFP fails to elicit any ISRE activation, unlike with infection with SeVAC (Fig. 5b). Although this phenomenon could be attributed solely to the inhibitory activity of viral IFN antagonist proteins, it is also plausible that infection with a wild-type-like virus simply does not produce RIG-I PAMPs in sufficient quantity to activate measurable ISRE activity. In agreement with this hypothesis is a recent study showing that, even in the absence of an IFN antagonist protein, most virus-infected cells still do not activate IFN production [37]. Indeed, during infections with SeV-GFP only minimal ISRE-RFP expression can be seen over the course of several days of infection (Fig. 5c). Cells positive for both red fluorescent protein (RFP) and GFP signal are very rare, indicating that ISRE activation in infected cells is a rare event requiring extensive virus replication (Fig. 5d).

Knocking down C protein leads to activation of innate immune responses

Identification of cbDI as a RIG-I ligand in infections with SeVAC led us to question how innate immune responses would be activated if we abrogated the function of the C protein with a small interfering (siRNA) knockdown approach post-infection. In this scenario, in contrast to initiating infections with a virus already rich in DIs, we would be infecting cells with a DI-free virus and subsequently removing the function of the C protein. Since the C ORF is expressed from the same mRNA as the P ORF, we had to take advantage of a recombinant virus which uncouples expression of P and C proteins. SeV-GFP-C contains a stop codon within the C ORF and a GFP-C protein fusion inserted between the M and F genes. Thus expression of C protein in this virus can be suppressed by siRNA targeting of the GFP ORF. Analysis of viral RNA from SeV-GFP-C stock revealed minimal presence of cbDIs, with overall levels similar to those in SeV-GFP stock, confirming that this virus, unlike SeVAC, does not favour accumulation of DIs during passaging (Fig. 4a). To assess the effect of C protein knockdown on the replication of the SeV-GFP-C and the immune response of the infected cells we transfected siRNA targeting GFP or control scrambled siRNA into HeLa cells and infected the cells 18 h post-transfection with SeV-GFP-C. Following infection we performed in-depth RNA-Seq analysis on both SeV and cellular RNA. Analysis of viral RNA sequences revealed that despite the loss of the C protein, the overall level of virus replication, as well as the relative levels of viral mRNAs, were remarkably similar between the siGFP and siCNTRL treatments (Fig. 6a). This is surprising, since the C protein has been shown to play a role in inhibition of virus replication [26]. We could see an approximately two-fold reduction in numbers of RNA sequences which mapped to the GFP mRNA; however the reduction in GFP protein expression was much more dramatic, likely indicating that the knockdown was occurring through translation suppression (Fig. 6b). In contrast to its minimal effect on virus replication, the loss of C protein had a dramatic effect on the cellular transcriptome. Despite similar levels of virus infection, we could see a clear activation of innate immune responses in siGFP-treated SeV-infected cells. Genome-wide transcriptome analysis revealed a great enrichment of innate immune-related gene cells, including IFNB1, confirming the established role of the C protein in suppressing innate immune responses (Fig. 6c). To determine whether cbDI accumulation was associated with the loss of the C protein we performed computational analysis on viral RNA from knockdown cells. We were unable to identify any reads associated with any cbDI RNAs in either siGFP- or siCNTRL-treated samples. This was not surprising, since the overall levels of viral RNA sequencing reads were much lower than those obtained from sequencing virus stocks, as cellular mRNA represented the majority of the RNA that was sequenced. To enrich for genomic viral RNA, as well as DI RNA, we performed pulldowns from siRNA-treated and SeV-GFP-C-infected cells with an antibody targeting SeV NP protein. Since all viral RNA genomes, including DI RNAs, are reported to be encapsidated within NP-covered ribonucleoprotein complexes, we postulated that isolating NP-bound RNA will likely result in the enrichment of viral genomic and DI RNAs. Sequencing profiles from NP pull-downs confirmed that we were enriching for genomic RNA species. However, we were again unable to identify any fusion RNA-Seq reads belonging to any cbDI molecules. It is possible that the DI RNAs generated during infection are...
present at too low a level to be detected with this approach. It is also possible that standard sequence preparation methods bias against isolation of DI RNA species, since they are optimized to detect longer RNAs, such as cellular mRNAs. The lack of isolation of any cbDI-containing reads made it impossible for us to conclude whether these RNAs were generated in C-protein-deficient cells and were associated with an activated immune response. In addition to utilizing the RNA-Seq approach, we attempted to detect cbDI RNA associated with SeVΔC through quantitative reverse transcription PCR (qRT-PCR). Although this assay is very effective at detecting cbDI RNAs when they are present in high amounts, it failed to do so in siGFP-treated samples, likely because the levels of cbDI are very low compared to full-length genomic RNA.

**Exogenous C protein limits levels of DI RNA**

The identification of DI RNA as a RIG-I ligand in SeVΔC infections was surprising, since this virus was propagated under DI-free growth conditions, such as plaque purification and low m.o.i. inoculation. Therefore we speculated whether the lack of the C protein inherently led to generation and/or accumulation of DIs. To this end, we measured the levels of DI1 during infection with SeVΔC in the presence or absence of exogenously expressed C.
protein. We saw a significant decrease in DI1 RNA levels, as well as genomic RNA, in the presence of C (Fig. 7). Since we could not generate a DI-rich stock from SeV-GFP, we postulated that perhaps the Cantell-strain C protein was deficient in its ability to limit DI replication. Analysis of C protein from both Z and Cantell strains, however, revealed no strain-specific differences in the ability of this protein to limit DI accumulation. The ability of exogenously expressed C protein to limit genome and DI RNA levels supports the hypothesis that DI RNA accumulation seen in the SeV ΔC mutant could be due to the lack of C function(s).

**DISCUSSION**

Activation of the RIG-I antiviral pathway by viral PAMPs represents an important early event in mounting an effective antiviral immune response. Although a large number of RNA molecules, both artificial and virus-generated, have been shown to be able to activate RIG-I signalling, the question of which RNAs activate RIG-I during natural infections is more complex. We previously showed that during infection with a Sendai virus rich in DI RNA the DI species functioned as the sole RIG-I PAMP, and we did not see any interaction between RIG-I and other viral RNAs. This left an open question: how does a DI-free virus activate a RIG-I-dependent immune response? In the current study we wanted to extend our work to a Sendai virus that is known to induce an IFN response by a different mechanism, through the absence of its IFN antagonist protein C. Surprisingly, we found that despite DI-free growth conditions, SeVΔC was generating high amounts of DI RNAs, and that these RNAs again served as exclusive ligands for RIG-I.

![Fig. 6. Knockdown of SeV C protein in infected cells. HeLa cells were transfected with siGFP or siCNTRL and infected with SeV-GFP-C (which expresses a GFP-C fusion protein). (a) Sendai RNA-Seq reads from siRNA-transfected and SeV- or mock-infected cells. The top two panels highlight very similar viral profiles in the presence or absence of C protein. (b) Comparison of loss of GFP RNA versus GFP protein, illustrating that the majority of the GFP protein is removed in siGFP cells. (c) RNA-Seq reads from siRNA-treated and infected cells were mapped to the human IFNB gene. Very strong IFNB expression can be seen in infected cells with GFP-C knockdown.](image-url)
during infection. Furthermore, analysis of a wild-type-like DI-free Sendai virus failed to detect any RIG-I-specific ligands. Therefore, it is possible that in the absence of DIs or other aberrant RNA generation, the RIG-I pathway may only be poorly activated. This phenomenon would fit well with the observation that only a very small portion of infected cells activate their IFN responses, which may represent those cells in which DIs or other abnormal virus replication products are generated [38–41].

The high accumulation of DI RNA in the absence of the SeV C protein is intriguing, and the mechanisms responsible for this phenotype are unknown. Interestingly, infection with SeV lacking C protein has previously been shown to lead to accumulation of dsRNA in infected cells, with concurrent PKR activation [42]. A similar observation has been made with a parainfluenza virus lacking the C ORF [43]. Since these studies did not specifically look for DI RNA in their experiments, it is possible that the detected dsRNA species belonged to the double-stranded regions found within the cbDI molecules. Identification of dsRNA in other paramyxoviruses may indicate that the ability of the C protein to limit DI levels is a more general phenomenon, and not limited exclusively to SeV. It is unclear whether the lack of normal C protein function results in a polymerase complex with a greater propensity to generate DIs, or if the C protein has a more direct role in controlling DI levels. It is also possible that C protein interaction with a host factor ultimately leads to control of DI generation or accumulation. A recent report examining SeV deficient in all four C ORF proteins 4C(−) concluded that the dsRNA species generated by this virus were not DIs, based on their lack of encapsidation and their ability to be detected by the J2-dsRNA-specific antibody versus the cpDI of SeV-Cantell, which cannot be detected by the J2 antibody [44]. Interestingly, in our studies we could not obtain a signal with J2 staining in either SeV-Cantell- or SeVΔC-infected cells, further supporting our conclusion that the IFN-inducing RNA species in SeVΔC-infected cells are DIs. The reason for the discrepancy between our study and that of Yoshida et al. [44] is unclear. The C proteins of these two SeVs were eliminated differently – mutation of the individual ribosomal start codons for 4C(−), and the inclusion of three consecutive stop codons within the C ORF shortly downstream of the Y2 start codon for deltaC – but this difference in itself does not explain the discrepancy.

Another recent study examining measles virus C protein reported accumulation of DI RNA in the absence of measles C, and furthermore the authors reported the accumulation of A-to-G editing, which is indicative of ADAR1 activity [45]. Examination of our own RNA-Seq data from SeVΔC-infected cells or viral stocks did not reveal any A-to-G editing events within the DI RNA sequences.

Our findings, as well as several other recent reports examining C-protein-deficient virus mutants, bring up the interesting possibility that mutant viruses which lack the ability to suppress the IFN response may do so not only through altered protein-mediated mechanisms, but also through accumulation of RNA products that function as cellular PAMPs [45, 46]. In fact, a number of reports indicate that manipulation of the viral genome may unintentionally lead to DI accumulation, irrespective of virus growth conditions [14, 47]. Therefore, it may be prudent to assess the make-up of viral RNA populations when examining phenotypes associated with a particular virus mutant.

The precise mechanisms responsible for DI generation remain unknown, but they are believed to be caused by errors in viral polymerase function. Historically associated with laboratory virus propagation conditions (e.g. high MOI infections), the existence of DI RNAs in natural infections remains a topic that has not been addressed in great detail. However, a number of recent studies indicate that these RNAs may be abundant in natural infections [48–52]. Advances in sequencing technology that allow global, in-depth analysis of viral populations are likely to reveal a great heterogeneity of viral RNA species, including DI RNAs, in infected hosts. The potential role that these RNAs may play in the activation of antiviral innate immunity in vivo is a topic of great interest and importance. This knowledge is especially critical for the design of effective vaccines and antivirals, whose efficacy often hinges on their ability to efficiently stimulate innate immune pathways. Currently, despite intensive investigation, we know very little about precisely how innate immunity is activated during virus infections.

---

Fig. 7. Exogenous C protein limits production of cbDI as well as genomic RNA. Relative expression of SeV cpDI or genomic RNA in the presence or absence of SeV C protein. SeV-Can or SeV-Z C protein-expressing plasmids were transfected into HeLa cells subsequently infected with either SeVΔC or SeV-Cantell. Viral RNA was measured at 24 h post-infection by q-RT-qPCR. Viral RNA levels in RFP-transfected control cells were set to 100%. Normalized data from three individual experiments are shown with SE.
**METHODS**

**Viruses**

All Sendai viruses with the exception of SeV-Cantell originated from SeV strain Z (SeV-Z). All SeV stocks were grown in the allantoic cavity of 10-day-old embryonated chicken eggs. Viruses were grown under conditions associated with minimal DI accumulation, i.e. inoculation at high dilution and plaque purification. Virus present in the allantoic fluid was analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining after virus pelleting. Virus titres were determined by plaque formation on LLC-MK2 cells. The construction and rescue of recombinant SeVΔC and SeV-GFP have been described previously [19, 36]. SeVΔC was generated by the inclusion of three consecutive stop codons downstream of the Y2 ATG. Therefore, this virus is lacking in expression of all four C proteins. This virus was rescued by three or four blind passages, with no evidence of viruses on Coomassie blue staining with purified virus from allantoic fluid. By passage six the virus reached a plateau at 10⁷ to 10⁸ p.f.u. ml⁻¹; passages six to eight were used for experiments. rSeV-GFP-C contains a GFP-C fusion insert between M and F ORFs, inserted within MluI sites in a backbone with a stop codon in endogenous C ORF.

**RIG-I IPs**

Endogenous RIG-I co-IPs were performed as previously described [11]. Briefly, 293 T cells were infected with viruses at an m.o.i. of 10 and the infections were allowed to proceed for 24 h. Cells were lysed in polysome lysis buffer and endogenous RIG-I/RNA complexes were isolated with an anti-RIG-I mouse monoclonal antibody generated by the Mount Sinai School of Medicine hybridoma facility. RIG-I-associated RNA was extracted with a phenol/chloroform extraction and ethanol precipitation. Co-IPs with exogenous RIG-I were performed by transfecting a plasmid expressing HA-tagged RIG-I protein (pCAGGS-HA-RIG-I) into 293 T cells 1 h post-infection. Cells were lysed 24 h post-infection and RIG-I/RNA complexes were isolated with an anti-HA antibody (H3663; Sigma). RNA was isolated by phenol/chloroform extraction and tested for immunostimulatory activity by transfection into 293 T ISRE-Fluc cells. Control IPs were performed with a mouse monoclonal IgG1 anti-GFP antibody (ab1218; Abcam). The efficiency of the IPs was verified through Western blot analysis.

**RNA deep sequencing and analysis**

Isolated RNA was prepared for deep sequencing as previously described [11]. Briefly, ribosomal RNA was depleted with a human Ribo-Zero kit (Epiconcept). The remaining RNA was then prepared into a sequencing library with an Illumina RNA-Seq kit, with the omission of the polyA-enrichment step. The resultant library was sequenced on the Illumina HiSeq platform at the Mount Sinai Genomics Center. Sequencing reads were processed and analysed with the assistance of the Mount Sinai genomics facility. RNA from virus stocks and from siRNA-treated SeV-infected cells was prepared for sequencing with the Illumina TruSeq Stranded Total RNA kit and sequenced as single-end 100nt reads on Illumina HiSeq at the Rockefeller Genomics facility. To identify cbDI RNA sequences from virus RNA-Seq data, we utilized open-source RNA-Seq analysis tools. The quality of the FastQ sequencing reads was initially assessed through FastQC and high-quality sequences were selected with the FASTQ Quality Trimmer. High-quality reads were aligned to either SeV-Z or SeV-Cantell genomes with TopHat2. Fusion reads were identified by utilizing the fusion search function within TopHat2. Viral alignments were performed with the following TopHat2 settings: max. edit distance – 4, final read mismatches – 4, max. insertion length – 3, max. deletion length – 3, maximum number of alignments – 20, minimum intron length – 5, maximum intron length – 15 000, number of mismatches allowed in each segment for reads mapped independently – 2, minimum length of read segments – 15, fusion search anchor length – 10, fusion search minimum distance – 5, fusion search read mismatches – 3, fusion search multireads – 5 and fusion search multipairs – 2.

**Knockdown of C protein in SeV-C-GFP-infected cells**

HeLa cells were transfected with an siRNA targeting the GFP ORF (Thermo Fisher AM4626) or scrambled control siRNA. At 18 h post-transfection the cells were either infected with SeV-C-GFP at an m.o.i. of 10 or mock infected. The knockdown efficiency was monitored by examining the levels of GFP expression in infected cells 24 h post-infection by fluorescence microscopy. Total RNA from infected cells was collected 24 h post-infection and prepared for RNA-Seq analysis with the Illumina TruSeq Stranded Total RNA kit. Sequencing data were analysed in a similar manner to that described above. Sequencing reads were aligned to both Sendai virus and human genomes with TopHat2. Human genome alignments were performed against 1000 genomes (project build 37;GRCh37). Individual gene expression levels were visualized using the Integrative Genomics Viewer (IGV). Reads mapping to the human genome were further subjected to differential expression analysis with HTSeq (Python) followed by EdgeR (R). Statistically significant (P<0.05) differentially expressed genes were analysed for functional enrichment using the NIH-based DAVID Functional Annotation tool (v 6.7). SeV NP IPs were performed with an anti-SeV NP antibody from siRNA-treated and SeV-C-GFP-infected cells in the same manner as was described for RIG-I IPs. NP-associated RNA was prepared for sequencing and analysed in the same manner as total RNA from SeV-C-GFP-infected cells.

**C protein complementation**

293 T cells were transfected with plasmids expressing SeV C protein (pEBS-SeV-C) [53] or a control (RFP-expressing) plasmid. At 24 h post-transfection, the cells were infected with SeVAC and infections were allowed to proceed for 24 h. At this time, total RNA was isolated from the infected cells and subjected to TaqMan quantitative PCR analysis for DII
RNA and genomic RNA levels. TaqMan analysis was performed with the following primer/probe combinations. DI detection: SeV-Z-dil-FWD CAAGACTTCCAGGTACAAA AGAG, SeV-Z-dil-REV ACGGCAGCAGGATTACACT and Sen-DI-hyb-probe (5’ FAM, 3’ quencher) CAAGATGGTA ACTGGGTCATTCE. SeV genome detection: Sen-gen-TQM-FWD GGGAGGAGGTGCTGTTATCC, Sen_gen_TQM-REV TAGCCGAGTCTCGTTGTA and Roche universal probe no. 82. One-step qRT-PCR was performed with Roche LightCycler 480 RNA Master Hydrolysis Probes (04991885001) kit and GAPDH housekeeping control multiplex assay (Roche 05190541001). Analysis was performed on the Roche LightCycler 480.

Funding information
This work was partially supported by NIAID grant U19AI083025 to A. G.-S.

Acknowledgements
We thank Richard Cadagan and Osman Lizardo for excellent technical assistance.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References


