Respiratory syncytial virus directly binds miR-26, miR-27, and let-7 to de-repress gene targets through canonical and non-canonical mechanisms

Sarah Ressel[1], Sujai Kumar[1], Jose Roberto Bermúdez-Barrientos[1], Katrina Gordon[1], Julia Lane[1], Jin Wu[2], Cei Abreu-Goodger[3], Jürgen Schwarze[4], Amy H. Buck[1,5,*]

[1] Institute of Immunology & Infection Research, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, UK

[2] Janssen Research & Development, Janssen Pharmaceutica NV, Turnhoutseweg 30, 2340, Beerse, Belgium.

[3] Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, UK

[4] Child Life and Health, Centre for Inflammation Research, University of Edinburgh, Edinburgh EH16 4TJ, UK

[5] Lead contact

*Correspondence: a.buck@ed.ac.uk

Abstract

Specific viruses can sequester or inhibit miRNAs through direct RNA-RNA interactions to enable their life cycle, but we still lack comprehensive understanding of the existence and functions of such interactions in respiratory virus infections. Here we optimise the CLEAR-CLIP protocol in A549 cells and show that Respiratory syncytial virus (RSV) directly binds miR-26, miR-27, and let-7. We show a significant global up-regulation of miR-26 targets upon RSV infection as well as up-regulation of miR-27 targets involved in cell cycle regulation and immune signalling. Through genome analysis of miRNA-target reads we further show that pseudogenes which overlap transcriptional regulatory regions are functional miRNA targets. We identify and validate an interaction between miR-27 and a pseudogene-derived lncRNA that overlaps an enhancer and show that this interaction regulates expression of *Carcinoembryonic antigen-related cell adhesion molecule 1* (*CEACAM1)*. This work demonstrates that RSV directly interacts with host miRNAs to de-regulate host gene expression and provides biochemical support for the ability of miRNAs to directly regulate gene transcription linked to immune signalling.

Keywords:

miRNA, respiratory syncytial virus, RSV, regulation of gene expression, pseudogenes, lncRNA, regulatory RNA, CLEAR-CLIP

INTRODUCTION

Viruses have evolved many tactics to manipulate host cells to facilitate their life cycles, but we still lack a comprehensive understanding of all the molecules and mechanisms involved. Over the last two decades, various studies have revealed the ability of specific viruses to interact with cellular miRNAs through direct RNA-RNA interactions. Some of the viral-miRNA interactions have been shown to increase genome stability, as first reported for hepatitis C virus (HCV) and miR-122 and more recently for miR-17 with members of the *Pestivirus* genus [1, 2, 3, 4]. Other viral-miRNA interactions inhibit the cellular miRNA by targeting it for degradation, as discovered in Herpesvirus saimiri (HVS) and murine cytomegalovirus (MCMV) that both target miR-27 with distinct viral transcripts [5, 6, 7, 8]. Viruses can also sequester host miRNAs through interactions with viral transcripts to de-repress cellular

miRNA targets without causing degradation of the miRNA [3, 4]. It is difficult to predict miRNA-virus interactions since canonical miRNA binding requires complementarity with only 6-8 nucleotides (nt, defined as the seed region [9]) and miRNA binding sites could therefore be predicted by chance in most viral transcripts.

Biochemical approaches have become an important complement to computational methods for miRNA target identification. Various methods have been developed to identify miRNA-mRNA interactions by UV crosslinking and immunoprecipitation (CLIP) of AGO proteins followed by an intermolecular ligation step to generate chimeric sequences of miRNA and its target, including CLASH and CLEAR-CLIP [10, 11, 12]. These methods have identified novel miRNA-target interactions in several viral families, in particular pestiviruses which interact with miR-17 and let-7 to regulate viral RNA translation and stability [4]. To date no miRNA-virus interactions have been directly identified in respiratory viruses, in part due to lack of high-quality data in the A549 cell line, a human lung cancer epithelial cell line that is routinely used to model gene regulation in the context of bacterial and respiratory virus infections as well as lung cancer, allergy, and asthma.

Here we adapted the CLEAR-CLIP protocol [12] for the A549 cell line in order to identify miRNA-target interactions that occur during infection by Respiratory syncytial virus (RSV) and determine whether this virus directly interacts with cellular miRNAs. RSV is a 15 kb negative-sense single-stranded RNA virus of the *Paramyxovirus* genus that has received a lot of attention recently due to surging cases after the corona virus pandemic. Despite the major clinical and economic burden in young children and the significant morbidity and mortality in the elderly caused by RSV, there is still no specific therapy and only limited active immunisation available [13]. There is currently no global biochemical data on the miRNA-target interactions that occur during RSV infection. Several studies have proposed a role of viral proteins in causing miRNA de-regulation during infection [14, 15, 16, 17]. However, it is not known whether RSV directly interacts with or regulates miRNAs through RNA-RNA interactions.

We developed a bioinformatic pipeline that maps the targets across the human and viral genomes, and accounts for interactions with elements beyond coding sequences (CDS) and 3' untranslated regions (3'UTRs). From this analysis we show that RSV interacts with the miRNAs miR-26, miR-27, and let-7. We further demonstrate that many miR-27 targets are de-repressed during infection including a large network of cell cycle genes. We further identify new targets of miR-27 associated with immune signalling including non-canonical targets derived from pseudogenes that overlap with regulatory regions. We validate a specific interaction of miR-27 with a long noncoding RNA (lncRNA) that derives from a pseudogene which overlaps a promoter flanking region (reannotated as enhancer in Ensembl release 109) downstream of *Carcinoembryonic antigen-related cell adhesion molecule 1* (*CEACAM1)* and show that this interaction regulates expression of CEACAM1. Our results demonstrate transcriptional regulation of CEACAM1 via interaction of miR-27 with a pseudogene-derived lncRNA. This work identifies new mechanisms by which miRNA de-regulation can impact immune signalling and adds RSV to a growing list of viruses that directly modulate miRNA activity through interaction with their genomes.

RESULTS

CLEAR-CLIP protocol optimisation in A549 cells

To generate a high quality CLEAR-CLIP dataset in A549 cells, we tested and modified several parameters compared to the protocols reported previously [4, 12, 18]. The first step of the protocol involves UV crosslinking to generate covalently linked protein-RNA complexes,

where different doses have been used across different cellular contexts [10, 12, 18, 19, 20]. We found a suitable dose in A549 cells to be 200 mJ/cm^2. In contrast, no crosslinking yielded fewer chimeric reads and was enriched for non-chimeric miRNA reads, as expected from previous literature [12]. Higher dose (400 mJ/cm^2) was associated with protein loss and fewer miRNA-target chimeras (Fig S1A-C, Table S1). Irradiated cells were immediately lysed and subsequently treated with RNases and the intermolecular ligation step was carried out with minor modifications (see Materials and Methods). A 3' adapter with an infrared dye allowed for visualisation of the RNA on SDS-PAGE (Fig S1D) and excision of the AGO-RNA complexes was carried out from the gel without the need for transfer to a membrane. We analysed two different molecular weight regions and found the high molecular weight AGO2 enriched for miRNA-target interactions (115-185 kDa, Fig S1C), consistent with previous reports [12, 21].

Replicate datasets were generated with A549 cells (1x10^8) either infected with RSV at a multiplicity of infection (MOI) of 0.1 or mock-infected for 48 h (for summary of protocol see Fig 1A). Chimeric reads were identified based on the requirement that the first part of the read maps to a mature miRNA. The remaining part of the read is then used to identify targets with a bioinformatic pipeline that maps reads to the human genome using Shortstack [22] and clusters mapped reads into target sites (Fig 1B, S1G). The percentage of chimeric reads ranged from 3-6 %, on par with previous reports in other cell types (Fig 1C). By mapping to the genome instead of the transcriptome our analysis can identify non-canonical target sites (beyond 3'UTRs or CDS). However, one challenge is the large number of overlapping genome annotations ("Multiple", Fig S1E). Here we classify "multiple" target sites that overlap 3'UTR as 3'UTRs for downstream analysis, since 3'UTRs are the best characterised and validated miRNA target categories, and since > 50 % of the "multiple" target sites overlapped 3'UTRs (Fig S1F). The remaining target sites with overlapping annotations were kept in the "multiple" category and not further classified. The initial analysis of seven datasets (n=3 in Mock and n=4 in RSV-infected) yielded a total of 389,168 and 220,938 unique target sites across Mock and RSV-infected samples, respectively. As a gauge for potential non-specific chimeras that could occur during sample preparation, we note that 10-18 % of the target sites were mapped to ribosomal RNAs (rRNAs, Fig 1D, Table S2); this is comparable to a similar percentage of rRNA chimeric reads in other CLEAR-CLIP reports: 12-17 % in Moore et al. [12] based on re-analysis of their data with our pipeline (Table S2).

Filtering chimeric reads for reproducibility, cluster length and seed sites enriches for miRNA-target interactions in 3' UTRs

Previous studies have shown that AGO scans for targets in a stepwise mechanism, searching first for complementarity with nucleotides 2-4 of the miRNA before extending complementarity to the nucleotides 5-8 [23]. To reduce the likelihood of non-specific and non-functional targets in our dataset, we required that chimeras were reproducible in ≥ 2 Mock or ≥ 2 RSV samples, and that the length of the cluster of target reads was < 150 nt (for reference clusters within ribosomal RNAs have a mean of >150 nt whereas those within 3'UTRs are <100 nt; Fig S1G, H). We then applied a further filter to require that target interactions contain canonical seed-based pairing (not allowing mismatches or bulges but allowing 6mer-offset as defined in Bartel [9]) since miRNAs are generally thought to require the interaction with the seed site for efficient inhibition of target translation [9]. Our filtering steps of reproducibility, length and seed site resulted in a subset of 7,539 and 2,898 high confidence (HC) target sites in Mock and RSV-infected samples respectively, with over 50% of the target sites mapping to 3'UTRs (Fig 1E). The HC target sites were enriched for

3'UTRs and CDS (compared to the unfiltered chimeric target list) and all the other target types were depleted, including rRNAs (Fig 1F).

miRNAs miR-26, miR-27 and let-7 interact with specific sites in the RSV genome but are not degraded upon infection

To examine whether RSV interacts directly with host miRNAs, we mapped non-chimeric (AGO binding) and chimeric reads from the CLEAR-CLIP datasets to the RSV genome. Chimeras and AGO binding sites were found across most RSV genes but were by comparison underrepresented in NS1 and L genes (Fig 2A, B). This appears to show selectivity as transcripts for NS1 and L genes are expressed at comparable levels based on RNA sequencing (Fig 2A, top panel). However, 89.1% of RSV-miRNA target sites are neither reproducible nor seed-mediated, suggesting scope for non-specific interactions (Table S1). Strikingly we found that nearly all seed-based chimeras with RSV derive from miR-26-5p, miR-27-3p and let-7-5p. These bind to specific positions in the M, SH, G and NS2 genes (Fig 2C).

Previous studies have shown that interactions with viral RNA can lead to miRNA degradation, in particular in the case of two different herpesviruses targeting miR-27 [5, 6, 7, 8] where the viral RNAs exhibit extensive pairing with the 3' end of the miRNA [24]. To determine whether miR-27, as well as miR-26 and let-7 were degraded during infection, we examined miRNA expression levels by small RNA-seq. The expression of all three miRNA families is slightly, but not significantly, upregulated in RSV infection at 48 hpi (hours post infection, Fig 2D). In addition, the interactions with RSV do not involve 3' end pairing (Fig S2A-C). However, miR-27 is shortened over the time course of RSV infection from 21 nt to 20 nt - which is not the case for miR-26 or let-7 family members (Fig S2D-F). The effect that this shortening would have on target interaction is not clear, but it has been proposed 3' end variations affect AGO2 binding as well as miRNA-mRNA interactions [25]. This could explain the lower percentage of HC chimeras that we observe for miR-27 but not miR-26 or let-7 in RSV infection (Fig 2E).

miR-27 high confidence targets include genes linked to cell cycle regulation and immune signalling

To determine whether the interaction with the viral RNAs could be associated with global de-repression of host miRNA targets upon infection, we compared gene expression in Mock and RSV-infected A549 cells by RNA-seq. We see significant upregulation of miR-26b HC 3'UTR targets and a more subtle de-regulation of miR-26a and miR-27a and miR-27b HC 3'UTR targets whereas there was no de-repression for let-7 targets (Fig S2G-I). Our analysis of RSV, combined with the previous literature in herpesviruses, suggests that miR-27 is directly functionally disrupted by both RNA and DNA viruses and may play important functions in the infection context. Although upregulation of miR-27 targets upon infection was not significant at a global level, we note that many previously validated targets of miR-27 were found to be upregulated upon infection, including WEE1, PLK2 and SPRY2 (Table S3) [26, 27, 28].

To probe this further we used the 396 HC 3'UTR targets of miR-27 identified in A549s here to characterize the pathways and networks targeted in human lung cells. GO term analysis using ShinyGO [29] identified several terms related to metabolic processes and cell cycle as enriched among miR-27 targets (Fig 3A), as well as PI3K-Akt and MAPK signalling when investigating KEGG pathways (Fig 3B)[30, 31].

To determine whether the HC 3'UTR targets were functionally regulated by miR-27, we overexpressed or inhibited miR-27 using synthetic mimics and inhibitors. Additionally, we

generated a lentiviral transduced A549 cell line to stably overexpress miR-27a. Subsequent total RNA-seq showed that overexpression of miR-27 resulted in downregulation of miR-27 HC 3'UTR targets compared to controls (Fig 3C, E). In contrast, miR-27 inhibitors caused a subtle but significant upregulation of miR-27 HC targets (Fig 3D). In all the conditions miR-27 HC 3'UTR targets (395 genes) showed a significantly stronger suppression (and upregulation for the miR-27 inhibitor) compared to analysis of all the miR-27 3'UTR targets for which at least 1 chimera could be found (4415 genes), highlighting the value of our filtering approach to identify functional targets. We observed stronger downregulation using miRNA mimics compared to the miR-27a overexpressing cell line which is most likely due to the difference in miR-27 levels (20-fold overexpression in mimics compared to 3-fold in the cell line; Fig 3F).

The miR-27 mimics significantly reduced the expression of 28 HC 3'UTR target genes associated with cell cycle regulation (Fig 3G, red/purple dots), some of which were also upregulated in RSV infection (labelled dots Fig 3G, Table S3). Although miR-27 has been associated with promotion of cell cycle progression by targeting negative regulators in different cancers [27, 32, 33, 34, 35] most studies have only focussed on a single target gene. The CLEAR-CLIP data uniquely demonstrates that miR-27 directly regulates many cell cycle-associated genes simultaneously, consistent with its role in promotion of cell cycle [27, 32, 33, 34, 35]. Our high confidence 3'UTR chimeras include previously validated targets such as WEE1, PLK2 and CCNG1 [26, 27, 36, 37], as well as many which have not been associated with miR-27 regulation before, including SGK1, TLK2 and MDM2 (Fig 3G, Table S3). These findings suggest that de-regulation of miR-27 activity by RSV could contribute to cell cycle arrest in infected cells which is known to occur and promote virus replication [38, 39].

In addition, miR-27 mimics also caused significant downregulation of target genes associated with PI3K-Akt and MAPK signalling (Fig 3G, blue/purple dots), some of which were also upregulated in RSV infection (Table S3). Interestingly, we found that one of the most strongly downregulated genes in response to miR-27 mimics is *CSF1* (*colony stimulating factor 1*, Fig 3G), a cytokine involved in intercellular signalling and macrophage activation [40]. CSF1 protein levels are also slightly reduced in the miR-27a overexpressing cell line (Fig 3H, I) further supporting regulation by miR-27. Our pathway analysis identified additional miR-27 targets associated with "response to cytokine" including Chemokine (C-X-C motif) ligand 2 (*CXCL2*) and *Mitogen-Activated Protein Kinase Kinase Kinase Kinase 3* (*MAP4K3*) (Table S4) which were also significantly downregulated in response to miR-27 mimics (Table S5). Collectively, these data uncover robust targeting of a network of 28 cell cycle genes by miR-27 and suggest a potential role of miR-27 in signalling to immune cells through regulation of genes such as *CSF1* and *CXCL2* which we identify here as targets of miR-27.

Non-canonical targets of AGO2:miR-27 suggest a role in transcriptional gene regulation

Recently, there has been increasing evidence that AGO:miRNAs can also regulate transcription through interaction with promoter-associated lncRNAs, leading to either activation or inhibition of gene expression in the nucleus [41, 42, 43, 44, 45]. We therefore examined if any classes of regulatory regions were enriched in our HC target sites. Regulatory regions that overlap pseudogenes (PSGs) were found to be the most enriched category in HC chimeras compared to total target datasets (Fig 4A, overlapping PSG). Pseudogene-derived lncRNAs are generally believed to act as miRNA sponges that interfere with regulation of the parental gene by a miRNA [46]. However, we hypothesised that some of the interactions might

play a role in AGO:miRNA-mediated regulation of gene transcription since many HC PSG target sites overlap transcriptional regulatory regions.

To test this hypothesis, we focused on the top target for miR-27 that overlapped a regulatory region and PSG (top target defined by total read counts and presence in all three Mock samples). This miR-27 target originated from a highly conserved promoter flanking region (reannotated as enhancer in Ensembl release 109) that overlaps *zinc finger, AN1-type domain 5* (*ZFAND5*) processed PSG on chromosome 19 (Fig 4B). The binding site is located in the CDS of the PSG as well as the parental gene (the PSG does not contain introns or UTRs, Fig S3A). One challenge with studying PSGs is the ability to map reads confidently to either PSG or the parental gene. Nonetheless, some unique mapping reads from RNA-seq confirmed that an RNA is expressed from the PSG locus (Fig 4B), which is further supported by a previous study analysing data from the Genotype-Tissue Expression (GTEx) project [47]. According to our data, the PSG is not a full-length transcript but instead a lncRNA which is missing part of exon 5 as well as exon 6 and 7 (Fig 4C, S3A). We will therefore refer to this RNA as PSG lncRNA.

We first examined whether the parental gene *ZFAND5* was regulated by miR-27. The sequences of parental gene and PSG of *ZFAND5* are not identical but have a few single nucleotide variations that enable them to be distinguished by RNA-seq (Fig S3A). Based on the RNA-seq analysis with miR-27 mimics and the A549-miR-27a overexpression cell line, we find no significant miR-27-mediated regulation of *ZFAND5* mRNA (Fig 4D, S3B), which was further confirmed by Western blot of ZFAND5 protein levels (Fig S3D). There was also no effect on the PSG lncRNA by miR-27 manipulation (Fig 4D, S3C). These data suggest that the interaction of miR-27 with the PSG lncRNA does not impact the PSG expression, nor does it impact *ZFAND5* gene expression (the PSG does not appear to act as a “sponge” for the miRNA). Since the PSG overlaps with a regulatory region, we therefore tested if the interaction of miR-27 with the PSG lncRNA could be involved in regulating gene expression of nearby genes. The closest gene was *CEACAM1* (Carcinoembryonic antigen-related cell adhesion molecule 1), an important regulator of immune cell activation [48] which we observe to be up-regulated upon RSV infection (Fig 4B). miR-27 overexpression using the stable cell line (A549 miR-27a) resulted in a reduction in CEACAM1 mRNA and protein levels (Fig 4D-G). The miR-27 mimic showed the same trend, while miR-27 inhibitor slightly increased *CEACAM1* expression (Fig 4E). Our data do not suggest direct regulation of *CEACAM1* by miR-27 because we do not observe HC 3’UTR chimeras with *CEACAM1* RNA. In addition, we note that *CEACAM1* does not have any predicted binding sites for miR-27 in the TargetScan 8.0 database [49]. To test if the regulation of *CEACAM1* by miR-27 requires the PSG lncRNA, we inhibited the PSG lncRNA using LNA GapmeRs (Qiagen, Fig 4H). We found that CEACAM1 mRNA and protein levels increased, in parallel with the reduction in PSG lncRNA (Fig 4I-K), similar to the trend we see when miR-27 is knocked down. Knockdown of the PSG lncRNA also partly restored CEACAM1 levels in miR-27a overexpressing cells (Fig S3J). Although the GapmeR also reduces ZFAND5 expression (Fig S3G-I), our data suggest that ZFAND5 itself does not regulate *CEACAM1* since siRNA knockdown of *ZFAND5* did not affect CEACAM1 mRNA or protein levels (or the levels of the PSG lncRNA, Fig 4J, K, S3E, F). These data suggest that both the lncRNA and miR-27 are needed for the suppression of *CEACAM1* expression. Consistent with transcriptional regulation of *CEACAM1* by PSG lncRNA and miR-27:AGO2, we confirmed nuclear localisation of AGO2 and miR-27 as well as nuclear enrichment of PSG lncRNA in A549 cells (compared to ZFAND5 gene, which is not enriched in the nucleus, Fig 5A-C). We also show that a small proportion of AGO2 is associated with the chromatin containing fraction (Nuc insoluble, Fig 5D).

Collectively, these data suggest that lncRNAs derived from regulatory regions can be involved in miRNA-mediated regulation of gene expression in the nucleus and present another function for pseudogenes that relates to the miRNA binding sites encoded within them.

DISCUSSION

Here we adapted CLEAR-CLIP as a powerful method to investigate genome-wide miRNA-target interactions in Mock- and RSV-infected A549 lung epithelial cells. After initial optimisation, we were able to recover a high number of miRNA-target chimeras in both conditions which allowed us to not only investigate the host gene target network of miRNAs in these cells but also interactions with viral RNAs. We show for the first time that RSV directly interacts with three host miRNA families through seed pairing: miR-26, miR-27, and let-7. Previous studies have shown that miR-27 is targeted and degraded by non-coding RNAs from HVS as well as MCMV [5, 6, 7, 8]. Notably, miR-27 has a subtle antiviral effect on MCMV [50] as well as different respiratory viruses *in vitro*, including multiple influenza strains, RSV A2 and a more pronounced effect against RSV Bt2a, a recent clinical isolate of RSV [51]. Additionally, a previous study in MCMV found that viral mutants unable to degrade miR-27 showed a significant reduction in viral titre in the acute infection in the lung [8].

Here we use the identification of 396 high confidence targets by CLEAR-CLIP combined with gene expression analysis of miR-27 modulated cells to systematically examine functional targets of miR-27 in lung epithelial cells. We show that the inhibition of miR-27 activity by RSV is associated with upregulation of pathways important for RSV infection and replication including PI3K/Akt and MAPK signalling as well as control of cell cycle [39, 51, 52, 53, 54, 55, 56, 57]. In addition, we identify as direct miR-27 targets the cytokine *CSF1* as well as the chemokine *CXCL2* which play important roles in intercellular communication [58]. This suggests that miR-27 target regulation not only impacts infected epithelial cells but could affect signalling to immune cells and other cells in the lung environment.

Using our method to select for high confidence target sites, we also show that pseudogenes (PSGs) overlapping regulatory regions are a strong candidate for functional miRNA interactions. We postulate that pseudogenes which contain miRNA binding sites in common with their parent genes can adopt novel functions in the pseudogene, beyond a role as miRNA sponges as previously described [46]. We provide evidence that an interaction between miR-27 and a lncRNA derived from a pseudogene can functionally regulate transcription of a nearby gene, *CEACAM1*. We show that both miR-27 and the ZFAND5 PSG lncRNA are required to suppress *CEACAM1* expression. This work interfaces with a growing body of literature demonstrating miRNAs and AGOs can operate in the nucleus to silence or activate gene expression and provides a natural example of this form of miRNA regulation in the context of RSV infection [59].

In this study we have focussed on miR-27 targets, however the interaction of RSV with miR-26 and let-7 are also intriguing and merit further attention. miR-26 has previously been reported to regulate antiviral immunity in several different viral infections [60, 61, 62, 63] and has been implicated in the inhibition of TLR4 signalling in RSV infection [64]. Here we demonstrate that RSV has evolved to directly inhibit the functional activity of this miRNA, re-enforcing a central role of miR-26 in antiviral signalling. In contrast to miR-26 regulation, the interaction of RSV with let-7 does not cause global de-regulation of let-7 targets but could be important during the viral life cycle; it is striking that pestiviruses also interact with this miRNA to enhance viral translation and stability [4].

In conclusion, this study adds RSV to a growing list of viruses known to directly interact with miRNAs during their life cycles and highlights pseudogenes as a class of miRNA targets that can mediate transcriptional regulation. As the field of RNA delivery advances, it will be timely to test the ability to interfere with virus-mediated miRNA de-regulation in this and other respiratory virus infections.

MATERIALS AND METHODS

Cell lines and culture conditions

A549 human lung epithelial cells (CCL-185, ATCC) and human epithelial type 2 (HEp-2, CCL-23, ATCC) cells were cultured in Dulbecco's modified eagle medium (DMEM, Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS, Gibco), 1 % L-Glutamine (L-glut, Gibco) and 1 % Penicillin/ Streptomycin (P/S, Gibco) at 37 °C and 5 % CO_2. Cell lines were regularly checked for mycoplasma contamination by PCR with primers (forward: 5'-GGGAGCAAACAGGATTAGATACCC-3' and reverse: 5'-TGCACCATCTGTCACTCTGTTAACCTC-3').

Virus and infection conditions

The human RSV, strain A2, used in this study was provided by Janssen Pharmaceuticals (Belgium). To visualise progress of viral infection a modified virus was used that had the eGFP gene introduced into the viral genome as described in Hallak et al. [65] and Guerrero-Plata et al. [66]. Briefly, in a series of minigenome steps the enhanced fluorescent protein (eGFP) gene flanked by RSV gene-start and gene-stop signals was inserted between the leader region and NS1 gene in a full-length cDNA copy of the viral genome. The plasmid was transfected into HEp-2 cells along with other plasmids encoding additional RSV proteins to support RSV transcription and replication. The virus was then passaged in HEp-2 cells. To determine viral titre, RSV immuno-plaque assays were performed in HEp-2 cells using 2-fold serial dilutions of the virus samples in DMEM. Cells were incubated with the virus dilution for 2 h before DMEM supplemented with 10 % FBS, 1 % L-glut, 1 % P/S was added, and the cells were incubated for further 22 h. The cells were then probed with 1:200 biotinylated RSV (Biorad) antibody and stained with 1:500 ExtrAvidin Peroxidase, with visualization of plaques by the addition of 3-Amino-9-EthylCarbazole (AEC) substrate. For infection assays in A549, the cells were seeded 24 h prior to infection and infected with RSV at a multiplicity of infection (MOI) of 0.1 in serum free DMEM (+1 % L-glut) for 1 h. Subsequently the cells were washed with serum free DMEM and incubated for a total of 48 h from the start of the infection in complete media (DMEM + 10 % FBS, 1 % L-glut) unless stated otherwise.

CLEAR-CLIP

The protocol described here was modified from the originally published CLEAR-CLIP[12] and modifications regarding buffers were guided by a variation of the protocol published by Bjerke and Yi [18]. We further adapted the protocol to include a custom 3' adapter from TriLink (developed here) and adaptation to remove need for membrane transfer (developed here). The modified protocol is described below.

<u>Sample preparation</u>

A549 cells were grown in 245 mm square dishes. One day post-seeding cells were at about 80-90% confluency and either RSV-infected at an MOI of 0.1 or Mock-infected for 48 h. The plates were then washed with PBS and UV crosslinked (245 nm) on ice, covered with ice-cold PBS using 200 mJ/cm^2 (or 400 mJ/cm^2/ not crosslinked when testing crosslinking intensities). The cells were harvested by scraping and immediately lysed in CLEAR-CLIP lysis buffer (50 mM Tris/HCl pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1 % Igepal, 0.5 % sodium deoxycholate, 0.1 % SDS, protease inhibitors (Roche)). After 10 min incubation on ice, lysates were DNase treated with 30 µL RQ1 RNase-free DNase and RNase treated with RNace-IT (Agilent) at a final concentration of 1:25,000 for 5 min at 37 °C with shaking. Lysates were cleared by centrifugation and 200 U RNasin RNase inhibitor (Promega) was added.

Immunoprecipitation

Protein G Dynabeads (Life Technologies) were prepared using 100 µL of beads coupled to 10 µg of AGO2 antibody (clone 11A9, Merck) or IgG control antibody (Rat IgG2a kappa Isotype Control (eBR2a), Invitrogen) per sample. Beads were equilibrated in lysis buffer before the lysate was added to the beads and incubated rotating at 4 °C overnight. The beads were then washed as followed.

- 2x CLEAR-CLIP lysis buffer (1x the beads were directly resuspended in buffer, 1x 3 min rotation at 4 °C)
- 3x HS-CLEAR (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1 % Igepal, 0.5 % sodium deoxycholate, 0.1 % SDS) for 3 min rotating at 4 °C each,
- 2x PNK+Tween (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2 % Tween-20)

All following reactions included 80 U RNase inhibitors and incubations were carried out with intermittent shaking.

5' phosphorylation, intermolecular ligation and 3' adapter ligation

The samples were then treated with 40 U of T4 polynucleotide kinase (PNK, 3' phosphatase minus, NEB) in a total reaction volume of 80 µL, including 10 mM ATP for 2.5 h at 20 °C to add a 5' phosphate group for the intermolecular ligation. Samples were subsequently washed 3x in PNK+Tween buffer. After the last wash, beads were flash centrifuged to remove residual buffer. The intermolecular RNA-RNA ligation was performed in a total volume of 100 µL using 62.5 U T4 RNA ligase 1 (NEB), 100 mM ATP, 15 % PEG-8000 at 16 °C overnight. The next day more ligase (25 U) and ATP (100 mM) were added to the reaction followed by further incubation for 5 h at 16 °C. The samples were washed 2x with CLEAR lysis buffer, followed by one wash in PNK/EDTA/EGTA buffer (50 mM Tris-HCl pH 7.4, 10 mM EDTA, 10 mM EGTA, 0.5 % Igepal) and two PNK+Tween washes. The samples were then treated with 8 U TSAP for 45 min at 20 °C and washed 2x in PNK+Tween buffer. A modified 3' sequencing adapter (3' CleanTag-IRDye800CW, TriLink, see Table S6, 1 µM) was ligated overnight at 16 °C in the dark using 800 U T4 RNA ligase 2 (truncated K227Q, NEB), 10 % PEG-8000 in a total reaction volume of 80 µL.

Elution from beads and SDS-PAGE purification of AGO2-RNA complexes

Beads were washed 2x with PNK+Tween and resuspended in elution buffer (15 µL 4x NuPAGE LDS sample buffer (Thermo Scientific), 45 µL PNK buffer (50 mM Tris-HCl pH 7.4, 50 mM NaCl, 10 mM MgCl2, 0.5 % TritonX-100, 5 mM β-Mercaptoethanol), 3 µL 1 M DTT). Samples were eluted from the beads at 70 °C with constant shaking. Eluate, containing AGO2-RNA complexes were run on a 4-12 % NuPAGE Bis-Tris protein gel for 150 min at 120 V in MOPS SDS running buffer (Thermo Scientific). Gels were run on ice to prevent RNA degradation and in the dark to protect the 3' adapter fluorophore. Fluorescent RNA was visualised using the Odyssey CLx scanner (Li-cor). The regions corresponding to 80-115 kDa or 115-185 kDa (see Table S1) were excised from the gel using a scalpel. Gel pieces were flash frozen on dry ice for 5 min and fragmented by centrifugation (6 min, 6,000 rpm) through a hole in a 0.5 mL tube into a 1.5 mL tube.

Protein digest, RNA extraction and sequencing

The eluted RNA-protein complexes were treated with 200 µg Proteinase K in 400 µL Proteinase K buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1 % TritonX-100, 10 mM Imidazole, 1 % SDS, 5 mM EDTA) and 1 µL RNasin RNase inhibitor (Promega). The mix was incubated for 2 h at 55 °C with constant shaking. To extract the RNA an equal volume of Acid-phenol:chloroform (Ambion) was added, and the extracted RNA was precipitated over the weekend. Precipitated RNA was washed with 70 % ethanol, air dried on ice and

resuspended in 2.5 µL RNase-free H_2O and 2.5 µL buffer 1 (CleanTag Small RNA library kit). The samples were then used for RNA library preparation using the CleanTag Small RNA library kit (TriLink) following manufacturer's instructions with the following modifications. As the 3' adapter was already ligated, library preparation was continued at the 5' adapter ligation step. The adapter was diluted 1:10, libraries were amplified using 20 PCR cycles and a size range of 160-238 bp was purified from the TBE gel (to avoid the miRNA only band). Samples were then sequenced by Edinburgh Genomics on Illumina NovaSeq using 100 single-end or by Genewiz (Azenta Life Sciences, Germany) using 150 nt paired-end as indicated in Table S1.

Transfection of miRNA mimics/inhibitors, siRNAs and GapmeRs

Synthetic mimic and inhibitors as well as controls were obtained from Dharmacon, siRNAs were obtained from ThermoFisher and GapmeRs were obtained from Qiagen (see Table S6). A549 cells were reverse transfected using 5 or 10 nM of the respective mimic/inhibitor/siRNA/GapmeRs in 0.3 % Lipofectamine 2000 (Invitrogen). These were pre-incubated with Lipofectamine in Opti-MEM I (Gibco) for 20 min and subsequently added to empty wells of either a 96 well or 6 well plate. The cell mix was added to the transfection mix in antibiotic free DMEM, supplemented with 10 % FBS and 1 % L-Glut. Transfected cells were incubated for 24 h at 37 °C and 5 % CO_2 before cell viability was measured using the CellTiter-Blue reagent (Promega) according to manufacturer's instructions and cells were infected.

Lentiviral transduction of A549 cells

A549 cells were transduced with lentiviral shMimics (Dharmacon) expressing miR-27a-3p together with turboRFP fluorescent reporter in the pSMART expression system under the human EF1α promoter. Cells were transduced at an MOI of 5 in serum free DMEM (+ 1 % L-Glut) with 8 µg/mL polybrene overnight. The next day FBS-containing media was added to adjust the serum concentration to 10 % and the cells were incubated for another two days. Then the cells were expanded and selected for successful transduction with 1.5 µg/mL puromycin. Serial dilutions of the polyclonal cell populations were then used to get single clones and monoclonal cell populations. Clones were assessed for RFP marker expression by microscopy and flow cytometry and miR-27 expression levels by reverse transcription quantitative PCR (see below).

RNA purification and sequencing

Total RNA was extracted from cells using miRNeasy mini kit (Qiagen) according to manufacturer's instructions. If RNA was being used for total RNA-seq, the samples were on-column DNase-treated according to manufacturer's instructions (Qiagen). RNA concentrations were determined using the Qubit RNA HS assay kit (Invitrogen). RNA quality was assessed using Bioanalyzer Pico chips (Agilent).

For total RNA-seq, libraries were prepared by the Genetics Core in Edinburgh using the NEBNext Ultra 2 Directional RNA library prep kit for Illumina and the NEBNext rRNA Depletion kit (Human/Mouse/Rat) according to the provided protocol. Sequencing was performed on the NextSeq 2000 platform (Illumina) using NextSeq 1000/2000 P3 Reagents (200 Cycles) v3. For small RNA-seq, 1 µg of RNA was treated with 5' polyphosphatase (Cambio) according to kit instructions, followed by ethanol precipitation of the RNA. Libraries were generated from 100 ng of RNA using the CleanTag Small RNA library kit (TriLink) according to the manufacture's instruction with the following modifications. Adapters were diluted 1:2, libraries were amplified using 15 PCR cycles and a size range of 147-238 bp

was purified from the TBE gel. Libraries were sequenced on the Illumina NovaSeq (Edinburgh Genomics) using 100-base singe end.

Reverse transcription and quantitative PCR

Extracted RNA was reverse transcribed using the QuantiTect RT kit (Qiagen, for total RNA) or miRCURY RT kit (Qiagen, for small RNAs) according to manufacturer's instructions. The resulting cDNA was diluted as per kit instructions and quantified with specific primers (Table S6) using either QuantiNova SYBR green assays (Qiagen, for cDNA from total RNA) or miRCURY SYBR green assays (Qiagen, for cDNA from small RNAs). Small RNA miRCURY qPCR primer assays were purchased from Qiagen. Primers for the lncRNA PSG were designed to match the lncRNA PSG perfectly but have a mismatch at the 3' end to the parental ZFAND5 gene. Mismatches located in the 3' end were shown impair PCR amplification [67, 68, 69] and allow for differentiation between lncRNA PSG and parental gene. ZFAND5 parental gene primers are located in exon 6 (forward primer) and exon 7 (reverse primer), which according to our data are not expressed at the PSG locus and should only amplify mature mRNA. Data were analysed using the $2^{-\Delta\Delta Ct}$ method [70].

Nuclear fractionation

The nuclear fractionation protocol was adapted from Burke and Sullivan[71]. Briefly, cells were harvested and washed with PBS before being lysed in CSKT buffer (10 mM PIPES (pH 6.8), 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 0.5 % Igepal, protease inhibitors (Roche)) on ice for 10 min. Nuclei were pelleted by centrifugation and cytoplasmic fraction (supernatant) was removed. Cytoplasmic fraction was either mixed with an equal amount of SDS lysis buffer (1 % SDS, 2 % β-Mercaptoethanol), boiled at 95°C for 10min and used for Western blot analysis or RNA was ethanol precipitated for subsequent RNA extraction with the miRNeasy mini kit (see above). Nuclei were washed in CSKT buffer and either lysed directly in SDS lysis buffer for Western blot analysis as described above or Qiazol for RNA extraction. To generate whole cell lysates, harvested cells were directly lysed in SDS lysis buffer/ Qiazol. Equal volumes were maintained for cytoplasmic and nuclear fractions. To test chromatin association of AGO2, the washed nuclei were lysed in CLEAR-CLIP lysis buffer for 20 min. Nuclear soluble fraction (supernatant) was separated from the insoluble fraction (pellet, containing chromatin) by centrifugation. The insoluble fraction was then sonicated 3x 10 s with 30 s on ice in between, followed by treatment with 30 µL DNase RQ1 for 5 min at 37 °C with constant shaking. Equal volumes of extracted RNA were used as input for RT-qPCR. Data were analysed as described in Gagnon et al.[72] with minor modifications. Briefly, data was normalised relative to GAPDH levels in each cellular compartment and the fold change was calculated compared to the whole cell lysate. As miR-27 RT-qPCR was performed with the miRCURY kit that is not designed for mRNA quantification, miR-27 levels were only normalised to the whole cell lysate.

Western blot

Cells were lysed in CLEAR-CLIP lysis buffer or processed as described for nuclear fractionation. Same volumes, equivalent to $1x10^5$ cells were mixed with 4x NuPAGE LDS sample buffer (Invitrogen) and run on a 4-12 % NuPAGE Bis-Tris protein gel in MOPS SDS running buffer (Invitrogen). Proteins were then transferred onto a PVDF membrane (Immobilon-FL, Merck) at 100 V for 105 min. Membranes were blocked with either 3 % milk powder or 5 % BSA in TBS + 0.1 % Tween-20 (TBS-T) and then incubated with the primary antibody overnight at 4 °C. Primary antibodies used in this study: AGO2 (1:1500, clone C34C6, Cell signalling, 2897), α-Tubulin (1:2000, Cusabio, CSB-PA004344), β-Actin (1:1000, Cell signalling, 4967), Calreticulin (1:2000, Cell signalling, 2891), CEACAM1

(1:20,000, Abcam, ab235598), CSF1 (1:1000 in BSA, Abcam, ab9693), Histone H3 (1:2000, Cell signalling, 4499), ZFAND5 (1:1000, Sigma, HPA018129). The next day the membrane was washed with TBS-T and incubated with the secondary antibody (Goat anti-Rabbit IgG (H+L) Secondary Antibody, DyLight 800, Thermo Scientific) diluted 1:20,000 in TBS-T containing 3 % milk or 5 % BSA for 1 h. Membranes were washed again before protein bands were visualised using the Odyssey CLx scanner (Li-cor). For quantification of protein levels, the bands were quantified using Fiji ImageJ [73] as outlined in Stael et al. Section 3.5 [74].

Northern blot

To detect RNAs of interest by Northern blot, 1 µM of the probes (Table S6) were 5' end labelled with [γ^{32}P] ATP (6000 Ci/mmol, Perkin Elmer) using the T4 polynucleotide kinase (10 U/µL, Thermo Scientific) and subsequently cleaned up with Microspin G-25 columns (GE Healthcare) according to the manufacturer's instructions. Equal amounts of extracted RNA (3.5 µg) were mixed 1:1 with Gel Loading Buffer II (Invitrogen) and loaded on a 15% polyacrylamide/8 M Urea gel. RNA Decade Marker (Thermo Scientific) was prepared according to the kit instructions, diluted in Gel Loading Buffer II, and run along the samples. The gel was run in 0.5x TBE at 120 V for 2 h. Samples were then transferred onto a Hybond-N membrane (Amersham Cat. RPN203N, 1 h at 80 V in 0.5x TBE) and chemically crosslinked at 55 °C for 2 h using a solution of 0.16 M EDC (N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride, Merck) with 0.13 M 1- methylimidazole in H2O (pH 8).[75] The membrane was washed once with 2x SSC (20x SSC, Fisher Scientific) and then pre-hybridised in 4 mL of Perfecthyb Plus hybridisation buffer (Sigma-Aldrich). 5' end labelled probes were heated to 90 °C for 30 s and hybridised to the membrane over night at 42 °C with rotation. The next day the membrane was washed once with the following pre-heated buffers: Wash buffer 1 (2x SSC, 0.1 % SDS), wash buffer 2 (1x SSC, 0.1 % SDS) and wash buffer 3 (0.1x SSC, 0.1 % SDS). To visualise the RNA, the membrane was incubated on a Storage Phosphor screen for 24 h before it was imaged on a Typhoon Scanner (GE Healthcare).

Bioinformatic analysis

Small RNA sequencing analysis

Reads were first quality checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and low quality 3' bases (quality below 20) were removed. The 3' small RNA adapter TGGAATTCTCGGGTGCCA was trimmed using Cutadapt (version 1.15)[76] and only reads containing the adapter and ≥ 18 nt were retained. miRNAs were identified using QuickMIRSeq [77] with standard settings. The QuickMIRSeq output file with miRNA counts was then uploaded on Degust web tool (Version 4.1.3, Powell, 2015, DOI: 10.5281/zenodo.3258933) for differential expression analysis using the Voom/Limma method.

CLEAR-CLIP

Initial processing (quality control and trimming) was performed as described above. For paired-end reads only the forward read was used in the analysis. Non-chimeric miRNA reads were identified by mapping to the mature Human miRNAs (miRbase V22) using bowtie2[78]. All reads were mapped to mature Human miRNAs (miRbase V22) using bowtie2 using local sensitive parameters (--sensitive-local), as this allows for partial alignment of potentially chimeric reads to the miRNA database. If the first part mapped to a miRNA and the remaining part was ≥ 18 nt and mapped to the human genome, this read was classified as miRNA-target-chimera. For the initial round of sequencing (NovaSeq 6000, 100-base, single

end - except for 06_Mock (CLEAR_200_UP), see Table S1) we mapped the target part of the read to all human protein coding transcripts (GRCh38.p7, Gencode) using bowtie2. For the final dataset (NovaSeq 6000, 150-base, paired-end, as well as 06_Mock (CLEAR_200_UP), see Table S1), the target part was aligned to the human genome (GRCh38.p7, Gencode) with bowtie2, then the aligned portion is extracted and removed if it matches a low complexity sequence as determined by dustmasker. Those target read portions that were kept were then mapped to the human genome using ShortStack [22] as follows: allowing 1 nt mismatches, requiring reads to have a minimum expression level of 0.5 reads per million (mincov 0.5 rpm), using unique mapping reads as guidance to position multimapping reads (mmap u), disabling mapping positioning when no unique mapping reads are available as guidance (ranmax 'none'), evaluating all possible sites for multimapping reads (bowtie_m 'all'), merging clusters that are closer than 10 nucleotides apart (pad 10). If a read mapped to multiple locations and/or genome features, nearby unique mapping reads are used to guide the read count distribution of multimapping reads. Genome features included in the analysis were 3' untranslated regions (UTRs), coding sequences (CDS), 5' UTRs, introns, regulatory regions (CTCF binding sites, transcription factor (TF) binding sites, enhancer regions, open chromatin regions, promoter-flanking regions and promoter, unprocessed and processed pseudogenes), tRNAs, lncRNAs, rRNAs and miRNAs. Overlaps between annotations and the chimeric cluster regions were made with bedtools [79]. We then applied a set of filters to identify high confidence targets in these interactions. Firstly, target sites needed to be present in ≥ 2 samples of the same condition (RSV-infected or mock-infected). Secondly, as the requirement of only a 1 nt overlap could lead to exceedingly long target sites with little overlap the target site length should be shorter than 150 nt and thirdly the target site must contain a strong miRNA seed site (no-mismatches). The code used for this part of the analysis is available on Zenodo (https://doi.org/10.5281/zenodo.7924850). Chimeric miRNA-RSV interactions were identified as described above, except the target part of the read was required to not map to the human genome and map to the RSV genome (KT992094). miRNA-target binding patterns were predicted using the IntaRNA web tool [80]. Raw data as well as a processed table with all miRNA-host target interactions have been deposited on GEO. Accession numbers are listed in the data availability section. Re-analysis of samples from the Moore et al. study[12] was performed by removing degenerate adapters to keep those samples comparable to our data which did not include degenerate adapters.

Total RNA-seq

Read quality and adapter trimming was made with trimmomatic 0.39 [81] using the TruSeq3 PE adapters. Paired-end reads were aligned to the human genome (GRCh38.p7, Gencode) with hisat2 2.2.1 [82]. Visualization tracks were made within R with the packages: GenomicAlignments 1.28.0 [83], GenomicRanges 1.44.0, Rsamtools 2.8.0 (DOI:10.18129/B9.bioc.Rsamtools), rtracklayer 1.52.0 [84]. Each bam file was read into R with the readGAlignmentPairs function without reading secondary alignments. The strand mode for the GAlignmentPairs object was set to 2. Reads were split by plus or negative strand and coverage was calculated separately; the coverage of the negative strand was multiplied by -1 to show reads below the line. Both coverages were then joined in a single object and saved as bedGraph files, which were subsequently compressed with igvtools 2.5.3 [85] to tdf format.

A custom GFF3 file was built which included gene annotations available through Gencode and regulatory regions provided by Ensembl (release 100), both sources of annotations were merged in R using rtracklayer 1.52.0. All regulatory regions that overlapped any Gencode gene annotations were removed, and the feature type was set to "exon" to allow regulatory

regions to be quantified in downstream processes. The number of reads mapping to genomic features were quantified with the featureCounts function implemented in Rsubread (version 2.6.4) R package [86] using the following non-default parameters: isPairedEnd = TRUE, isGTFAnnotationFile = TRUE, fraction=TRUE, largestOverlap=TRUE, strandSpecific = 2.

The upset plot for multiple overlapping annotations was made with the ComplexUpset 1.3.5 R package. Degust web tool (Version 4.1.3, Powell, 2015, DOI: 10.5281/zenodo.3258933) was used for differential expression analysis using the Voom/Limma method, with min gene read count of 4 and min gene CPM of 1 in at least 1 sample. CDF plots were made with ggpubr R package. Target genes with more than one binding site were collapsed to ensure each gene was only present once per category.

Statistical tests

Details regarding the statistical tests used and the represented data can be found in the figure legends, or, for the differential expression analysis in the “Small RNA sequencing”/ “Total RNA-seq” section of the Bioinformatic analysis. Statistical tests were considered significant when P values were less than 0.05. Statistical analyses were performed via GraphPad Prism (Prism Software) or, for the Cumulative Distribution Function (CDF) plots in RStudio using the Wilcox.test function (alternative = “less” ("greater” for inhibitor data)). For the statistic comparison of HC targets vs all targets, HC targets were removed from the all targets data set.

Data Availability

The datasets and computer code produced in this study are available in the following databases:

- CLEAR-CLIP raw files and processed data table containing all miRNA-host target interaction: Gene Expression Omnibus accession number GSE232686 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232686)
- total RNA-seq of Mock and RSV-infected A549: Gene Expression Omnibus accession number GSE231788 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE231788)
- total RNA-seq of A549, A549 miR-27a, miR-27 mimics/inhibitors: Gene Expression Omnibus accession number GSE231787 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE231787)
- small RNA-seq of Mock and RSV-infected A549: Gene Expression Omnibus accession number GSE231784 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE231784)
- The original code used for the CLEAR-CLIP analysis is available at https://doi.org/10.5281/zenodo.7924850

Acknowledgements

This work was supported by the Wellcome Trust RCDF 201083/Z/16/Z and ISSF to A.H.B and Darwin Trust of Edinburgh for S.R. as well as collaborative research funding from Janssen Pharmaceuticals, Inc. (Beerse, BE). The authors thank Dr Alfonso Garrido-Lecca from the Yi group at the University of Boulder for sharing advice on modifications to the CLEAR-CLIP protocol. The authors thank the Macias group from the University of Edinburgh for sharing RT-qPCR primers, as well as the IgG control antibody. The authors also thank Dr Pablo Tristán-Ramos from the Heras group at the University of Granada for discussions around crosslinking conditions during the CLEAR-CLIP protocol optimisation and Dr Anton McCaffrey and Sabrina Shore from Trilink for discussions around the modified 3' adapter. Total RNA-seq library preparation and sequencing was carried out by the Genetics core at the Edinburgh Clinical Research Facility. Small RNA sequencing was carried out by Edinburgh Genomics at the University of Edinburgh.

Author Contributions

S.R. and A.H.B designed the study. All authors participated in the discussion and interpretation of the results. S.R. and J.L. performed the experiments. K.G. optimised the CLEAR-CLIP protocol for the fluorescent 3' adapter. J.W. provided the virus and reagents. S.K. developed the CLEAR-CLIP bioinformatic pipeline in communication with S.R. and C.A.G. J.R.B.B. modified the CLEAR-CLIP pipeline. S.K., J.R.B.B., SR performed the bioinformatic analysis of the (small) RNA-seq data. J.R.B.B and S.R. analysed the data with input from A.H.B and C.A.G. S.R. and A.H.B. wrote the manuscript. All authors revised the manuscript.

Declaration of Interests

Jin Wu is a paid employee of Janssen Research & Development, Janssen Pharmaceutica NV, Turnhoutseweg 30, 2340, Beerse, Belgium.

FIGURE LEGENDS

Figure 1 - miRNA target sites in RSV-infected A549 cells discovered by CLEAR-CLIP

A Overview of CLEAR-CLIP protocol.

B Schematic of bioinformatics pipeline: chimeric reads were defined as the first part of the read mapping to a mature miRNA with the remaining part ≥ 18 nt. Target sites were extracted and mapped to the human genome and subsequently clustered into target sites per miRNA. To remove background of non-specific interaction, target sites were filtered for reproducibility, length <150nt and presence of the miRNA seed.

C Percentage of chimeric vs non-chimeric reads mapping to the human genome in our samples (three biological replicates for Mock and four for RSV) compared to the Moore et al. study[12].

D Percentage of unique target sites across different RNA categories. Target regions are 3' untranslated region (3'UTR), coding sequence (CDS), 5'UTR, introns, regulatory regions (Regulatory), transfer RNA (tRNA), long noncoding RNA (lncRNA), ribosomal RNA (rRNA), miRNA, pseudogenes (PSG), multiple annotations (Multiple) or uncategorised (Uncat).

E Percentage of unique high confidence (HC) target sites across different RNA categories described in D).

F Enrichment of HC target sites compared to non-filtered target sites. Red asterisks indicate significant enrichment (One Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test performed separately for Mock and RSV samples, * $P \leq 0.05$, ** $P \leq 0.01$).

Figure 2 - miRNA interactions with the RSV genome do not lead to miRNA degradation but affect target interactions

A, B, C Genome browser view of the RSV genome (KT992094) showing total RNA-seq read coverage (top panel), as well as AGO2 and IgG binding (determined from non-chimeric reads mapping to RSV, lower panel) (A), all miRNA chimeras (grey) and seed-based, reproducible miRNA chimeras (blue) (B), and reproducible, seed-based binding sites for miR-26 (yellow), miR-27 (green), and let-7 (purple) (C).

D miRNA expression determined by small RNA-seq in counts per million (CPM) at 48 hours post infection (hpi) in Mock vs RSV-infected cells for miR-27a/b, miR-26a/b and let-7a/f. Shown are three biological replicates with mean ± SD. Significance was tested using multiple unpaired t-tests for each miRNA family member: all changes between Mock and RSV are non-significant.

E Percentage of host HC miRNA-target site reads of total chimeric reads in Mock vs RSV-infected samples for miR-27a/b, miR-26a/b and let-7a/f. Significance was tested using multiple unpaired t-tests (ns = non-significant, * $P \leq 0.05$). Different miRNA family members are highlighted by filled (miR-27a, miR-26a, let-7a) or open symbols (miR-27b, miR-26b, let-7f).

Figure 3 - miR-27 regulates genes associated with cell cycle progression and immune signalling

A, B Pathway analysis of miR-27 HC 3'UTR target genes using ShinyGO V0.77 [29] showing significant pathways (limited to 20 pathways, FDR < 0.05, pathway size: min. 2, max. 2000) for GO term Biological process (A) and KEGG pathways [30, 31] (B).

C, D, E Cumulative distribution function (CDF) plot of RNA-seq data comparing the $\log_2$ fold change of miR-27 HC 3'UTR target genes (red) vs all miR-27 3'UTR target genes (blue) and no target genes (black) between A549 cells transfected with 10 nM miR-27 mimics or control mimic (C), miR-27 inhibitor and control inhibitor transfected cells (D) and lentiviral transduced cell line stably overexpressing miR-27a (A549 miR-27a) and A549 wildtype (A549 WT) cells (E).

F qPCR results of fold change in miR-27 levels in A549 WT, A549 miR-27a, A549 transduced with a non-targeting control (A549 NTC), miR-27 mimic transfected and control mimic transfected cells. Shown are biological replicates with mean ± SD. Significance was tested using One-way ANOVA with Welch's correction (* $P \leq 0.05$, ** $P \leq 0.01$).

G Volcano plot of RNA-seq data showing $\log_2$ fold change (Log2FC) in gene expression in miR-27 mimic transfected cells compared to mimic control transfected cells. Highlighted are: miR-27 HC targets (dark grey), miR-27 HC targets associated with cell cycle (red) or MAPK/PI3K-Akt signalling (blue). Purple dots highlight HC targets belonging to both pathway categories. Gene names are shown for genes that were upregulated ($\log_2$ fold change > 0) in the RNA-seq data of RSV vs Mock-infected cells.

H Western blot analysis of CSF1 protein expression in A549 WT or miR-27a overexpressing cells at 48hpi compared to β-Actin loading control.

I) Quantification of CSF1 protein levels from H) normalised to β-Actin, shown as mean ± SD for two biological replicates.

Figure 4 - Non-canonical targets of Ago2:miR-27 suggest a role in transcriptional regulation of *CEACAM1*

A Enrichment of HC target sites in regulatory regions compared to non-filtered target sites. Target sites were filtered to remove target sites from 3'UTRs and CDS. Shown are mean ± SD for three Mock and four RSV CLEAR-CLIP replicates combined. Red asterisks indicate significant enrichment (One Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test. Due to the low number of target sites in some of the categories, Mock and RSV samples were analysed together for more statistical power, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$). "Overlapping" indicates regulatory regions overlapping other genome annotations (pseudogenes (PSG), lncRNAs and introns).

B, C Genome browser view of the genomic location of top miR-27 regulatory target overlapping a PSG, showing the conservation using PhyloP100, features of the genomic locus, total RNA-seq coverage in Mock and RSV-infected samples (n = 3), as well as miR-27 target site coverage (chimeric miR-27) in Mock and RSV-infected CLEAR-CLIP samples for miR-27 binding site and nearby genome features (B) or a zoom into the promoter flanking/PSG region (C).

D RNA-seq data comparing gene/lncRNA expression in CPM in A549 WT cells and A549 cells overexpressing miR-27a (A549 miR-27a). Shown are three biological replicates with mean ± SD. Significance was tested with two-way ANOVA with Sidak's multiple comparisons test (ns = non-significant, ** $P \leq 0.01$).

E qPCR results of fold change in *CEACAM1* mRNA expression in response to miR-27 overexpression (A549 miR-27a, miR-27 Mimic) or inhibition (miR-27 Inhibitor) and controls (A549 NTC, Control Mimic, Control Inhibitor). Mimics/inhibitors were transfected at a final concentration of 10 nM. Data is shown as mean ± SD. Significance was tested with two-way ANOVA with Sidak's multiple comparisons test (ns = non-significant, ** $P \leq 0.01$).

F Western blot analysis of CEACAM1 protein at 24 hpi and 48 hpi (RSV MOI 0.1) in response to miR-27a overexpression compared to β-Actin loading control.

G Quantification of CEACAM1 protein levels from F) corrected for β-Actin and normalised to A549 WT. Shown are two biological replicates with mean ± SD. Significance was tested with two-way ANOVA with Sidak's multiple comparisons test (*P ≤ 0.05, ** P ≤ 0.01).

H, I qPCR results of lncRNA PSG expression (H) or *CEACAM1* expression (I) after treatment with 5 nM GapmeR targeting lncRNA PSG (anti-lncRNA PSG) or control. Shown are two biological replicates with mean ± SD. Significance was tested with unpaired two-tailed t-test ((ns = non-significant, * P ≤ 0.05)).

J Western blot analysis of CEACAM1 after treatment with anti-lncRNA PSG GapmeR, siRNAs against ZFAND5 (siZFAND5) and CEACAM1 (siCEACAM1) and controls in comparison to β-Actin loading control.

K Quantification of CEACAM1 protein levels from J) corrected for β-Actin and normalised to Untreated control. Shown are two biological replicates with mean ± SD. Significance was tested with One Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test.

Figure 5 - AGO2, miR-27 as well as the lncRNA PSG are present in the nucleus of A549 cells

A qPCR results for Malat1, RPL30, ZFAND5, PSG lncRNA and miR-27 (left to right) levels in cytoplasmic and nuclear fractions in Mock and RSV-infected cells. Data is shown as mean ± SD for two biological replicates.

B Western blot analysis of AGO2 protein levels in subcellular fractions showing whole cell lysate (WCL), cytoplasmic (Cyto), as well as nuclear (Nuc) fractions for Mock and RSV-infected cells. α-Tubulin and Calreticulin were used as cytoplasmic markers and Histone H3 as nuclear marker.

C Quantification of AGO2 protein levels in B). Data is shown as mean ± SD for four biological replicates and normalised to WCL Mock.

D Western blot analysis of cellular fractionation showing cytoplasmic fraction (Cyto), as well as the soluble and insoluble nuclear fractions (Nuc). α-Tubulin and Calreticulin were used as cytoplasmic markers and Histone H3 as nuclear marker associated with the chromatin.

Supplemental Figures

Figure S1 - CLEAR-CLIP protocol optimisation and multiple annotations

A UV crosslinking influences AGO2 protein recovery. Western blot analysis of protein lysates after crosslinking A549 cells with different doses, probed for AGO2 and α-Tubulin (loading control). Equivalent of 5 x 10^5 cells loaded per lane.

B UV crosslinking affects percentage of miRNA-target chimeras. Percentage of chimeric reads and miRNA only reads was calculated from total reads mapped to the human genome for the different crosslinking dosages.

C Gel purification can enrich for chimeric reads. Ratio of chimeric reads compared to miRNA only reads of two different molecular weight ranges purified from the SDS-PAGE (80-115 kDa vs 185-115 kDa), separated by different crosslinking conditions.

D Visualisation of infrared 3'adapter ligated RNA (left) and AGO2 protein (right) on the SDS-PAGE before purifying complexes from 115-185 kDa.

E Many target sites are mapped to genome locations with multiple annotations (Multiple). Percentage of unique high confidence (HC) target sites across different RNA categories.

F Upset plot of target sites with multiple annotations show that many target sites overlap 3'UTRs, regulatory regions, and introns. Intersections filtered for >150 counts.

G Schematic of the generation of target site clusters. Visualising that the required minimum overlap of 1nt can result in extensive target sites with only little overlap between reads.

H Density plot showing length distribution of target sites in nucleotides (nt). Dashed line indicates mean target site length. Left panel: all reproducible target sites; right panel: comparing target sites from 3'UTRs to rRNA target sites.

Figure S2 - miRNA length and de-repression of miRNA targets in RSV infection

A, B, C Predicted binding pattern of miRNA family members with RSV RNA, determined with IntaRNA [80] for miR-26 (A), miR-27 (B), let-7 (C).

D, F miRNA length in nt in Mock and RSV-infected samples at 48 hpi determined from small RNA-seq data for miR27 (D) and miR-26 and let-7 (F). Data is shown as mean ± SD of three biological replicates per condition.

E Northern blot showing miR-27 length over time course of RSV infection compared to loading control RNU6.

G, H, I Cumulative distribution function (CDF) plot of RNA-seq data showing the $\log_2$ fold change in expression of miRNA HC target genes vs non-target genes in RSV vs Mock-infected cells for miR-26 (G), miR-27 (H), and let-7 (I) family members. Wilcoxon test p values between miR-26a/b HC targets and no target are shown (ns = non-significant).

Figure S3 - miR-27 mediated regulation of CEACAM1 is not mediated through ZFAND5 parental gene regulation

A Schematic of the parental *ZFAND5* gene locus, as well as mature *ZFAND5* mRNA, ZFAND5 PSG and lncRNA PSG (top). Clustal Omega multiple sequence alignment of ZFAND5 CDS, ZFAND5 PSG and lncRNA PSG (bottom). Nucleotide differences are highlighted.

B, C Expression of parental *ZFAND5* mRNA (B) or lncRNA PSG (C) in A549 cells transfected with 10 nM miR-27 mimics or controls determined from total RNA-seq in CPM. Data is shown as mean ± SD for three (two for mimic control) biological replicates. Significance was tested with two-way ANOVA with Sidak's multiple comparisons test (ns = non-significant).

D Western blot analysis of ZFAND5 protein expression in A549 cells transfected with 10 nM miR-27 mimics/inhibitors or controls as well as A549 cells overexpressing miR-27a. Quantification of two (mimics) or three (overexpressing cell line) biological replicates is shown as mean ± SD below the blot.

E, F, G qPCR results of RNA expression for lncRNA PSG (E), *CEACAM1* (F) or *ZFAND5* (G) after treatment with 5 nM GapmeR (anti-lncRNA PSG and GapmeR control) or 10 nM siRNAs and siRNA control. Data is shown as mean ± SD for two biological replicates. Dashed line indicates a fold change of 1 compared to Untreated control.

H Western blot analysis of ZFAND5 protein expression after treatment with 5 nM GapmeR (anti-lncRNA PSG, GapmeR control) or 10 nM siRNAs and control compared to β-Actin loading control.

I Quantification of ZFAND5 protein expression in H). Data is shown as mean ± SD for two biological replicates. Significance was tested with One Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test.

J Western blot analysis of CEACAM1 after treatment with 5 nM anti-lncRNA PSG GapmeRs and controls in A549 WT and miR-27a overexpressing cells compared to β-Actin loading control.

Supplemental Tables

Table S1: Summary of CLEAR-CLIP data sets generated in this study, as well as CLEAR-CLIP samples generated by Moore et al. Information for each sample includes sample name, UV intensity used for crosslinking (UV), molecular weight of protein-RNA complexes purified, information about sequencing run, as well as information about non-chimeric and chimeric reads. ND = not determined.

Table S2: Percentage of unique target sites in different RNA categories comparing our samples to the CLEAR-CLIP samples from Moore et al.[12].

Table S3: Differentially expressed genes in RSV vs Mock-infected cells at 48 hpi, MOI of 0.1 (log2FC = $\log_2$ fold change, FDR = False discovery rate, miR-27ab_HC = high confidence target of miR-27a/b with 1 = miR-27 HC target & 0 = non-miR-27 HC target).

Table S4: ShinyGO Biological process pathway analysis (nGenes = number of genes).

Table S5: Differential gene expression in response to miR-27 mimics/inhibitors transfection or A549 cells overexpressing miR-27a (MI = mimic, MIC = mimic control, INH = inhibitor, INHC = inhibitor control, 27a = A549 overexpressing miR-27a, NTC = A549 cells overexpressing non-targeting control, WT = A549 wildtype, log2FC = $\log_2$ fold change, AdjPval = adjusted P value).

Table S6: Oligonucleotides, miRNA mimics & inhibitors, siRNAs and GapmeRs used in this study.

882 **REFERENCES**

883 1. Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of Hepatitis C Virus RNA
884 Abundance by a Liver-Specific MicroRNA. *Science* **309**, 1577-1581 (2005).

885
886 2. Shimakami T*, et al.* Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc*
887 *Natl Acad Sci U S A* **109**, 941-946 (2012).

888
889 3. Luna JM*, et al.* Hepatitis C virus RNA functionally sequesters miR-122. *Cell* **160**, 1099-1110
890 (2015).

891
892 4. Scheel TK*, et al.* A Broad RNA Virus Survey Reveals Both miRNA Dependence and Functional
893 Sequestration. *Cell Host Microbe* **19**, 409-423 (2016).

894
895 5. Cazalla D, Yario T, Steitz JA. Down-regulation of a host microRNA by a Herpesvirus saimiri
896 noncoding RNA. *Science* **328**, 1563-1566 (2010).

897
898 6. Buck AH*, et al.* Post-transcriptional regulation of miR-27 in murine cytomegalovirus
899 infection. *RNA* **16**, 307-315 (2010).

900
901 7. Libri V*, et al.* Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target. *Proc*
902 *Natl Acad Sci U S A* **109**, 279-284 (2012).

903
904 8. Marcinowski L*, et al.* Degradation of cellular mir-27 by a novel, highly abundant viral
905 transcript is important for efficient virus replication in vivo. *PLoS Pathog* **8**, e1002510 (2012).

906
907 9. Bartel DP. Metazoan MicroRNAs. *Cell* **173**, 20-51 (2018).

908
909 10. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by
910 CLASH reveals frequent noncanonical binding. *Cell* **153**, 654-665 (2013).

911
912 11. Grosswendt S*, et al.* Unambiguous identification of miRNA:target site interactions by
913 different types of ligation reactions. *Mol Cell* **54**, 1042-1054 (2014).

914
915 12. Moore MJ*, et al.* miRNA-target chimeras reveal miRNA 3'-end pairing as a major determinant
916 of Argonaute target specificity. *Nat Commun* **6**, 8864 (2015).

917
918 13. Mammas IN*, et al.* Update on current views and advances on RSV infection (Review). *Int J*
919 *Mol Med* **46**, 509-520 (2020).

920
921 14. Bakre A*, et al.* Respiratory syncytial virus modifies microRNAs regulating host genes that
922 affect virus replication. *J Gen Virol* **93**, 2346-2356 (2012).

923

924 15. Thornburg NJ, Hayward SL, Crowe JE, Jr. Respiratory syncytial virus regulates human
925 microRNAs by using mechanisms involving beta interferon and NF-kappaB. *mBio* **3**, (2012).

926
927 16. Bakre A, Wu W, Hiscox J, Spann K, Teng MN, Tripp RA. Human respiratory syncytial virus
928 non-structural protein NS1 modifies miR-24 expression via transforming growth factor-beta.
929 *J Gen Virol* **96**, 3179-3191 (2015).

930
931 17. Bakre AA, Harcourt JL, Haynes LM, Anderson LJ, Tripp RA. The Central Conserved Region
932 (CCR) of Respiratory Syncytial Virus (RSV) G Protein Modulates Host miRNA Expression and
933 Alters the Cellular Response to Infection. *Vaccines (Basel)* **5**, (2017).

934
935 18. Bjerke GA, Yi R. Integrated analysis of directly captured microRNA targets reveals the impact
936 of microRNAs on mammalian transcriptome. *Rna* **26**, 306-323 (2020).

937
938 19. Fu X, Liu P, Dimopoulos G, Zhu J. Dynamic miRNA-mRNA interactions coordinate gene
939 expression in adult Anopheles gambiae. *PLoS Genet* **16**, e1008765 (2020).

940
941 20. Stebel S, Breuer J, Rossbach O. Studying miRNA-mRNA Interactions: An Optimized CLIP-
942 Protocol for Endogenous Ago2-Protein (2022).

943
944 21. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA
945 interaction maps. *Nature* **460**, 479-486 (2009).

946
947 22. Axtell MJ. ShortStack: comprehensive annotation and quantification of small RNA genes.
948 *RNA* **19**, 740-751 (2013).

949
950 23. Chandradoss Stanley D, Schirle Nicole T, Szczepaniak M, MacRae Ian J, Joo C. A Dynamic
951 Search Process Underlies MicroRNA Targeting. *Cell* **162**, 96-107 (2015).

952
953 24. McCaskill J, Praihirunkit P, Sharp PM, Buck AH. RNA-mediated degradation of microRNAs: A
954 widespread viral strategy? *RNA Biol* **12**, 579-585 (2015).

955
956 25. Tomasello L, Distefano R, Nigita G, Croce CM. The MicroRNA Family Gets Wider: The IsomiRs
957 Classification and Role. *Frontiers in Cell and Developmental Biology* **9**, (2021).

958
959 26. Wuchty S*, et al.* Prediction of Associations between microRNAs and Gene Expression in
960 Glioma Biology. *PLOS ONE* **6**, e14681 (2011).

961
962 27. Tian Y*, et al.* MicroRNA-27a promotes proliferation and suppresses apoptosis by targeting
963 PLK2in laryngeal carcinoma. *BMC Cancer* **14**, 678 (2014).

964
965 28. Liu C*, et al.* MicroRNA-27b inhibits Spry2 expression and promotes cell invasion in glioma
966 U251 cells. *Oncol Lett* **9**, 1393-1397 (2015).

967
968 29. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants.
969 *Bioinformatics* **36**, 2628-2629 (2019).

970
971 30. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating
972 viruses and cellular organisms. *Nucleic Acids Research* **49**, D545-D551 (2020).

973
974 31. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data
975 integration and visualization. *Bioinformatics* **29**, 1830-1831 (2013).

976
977 32. Mertens-Talcott SU, Chintharlapalli S, Li X, Safe S. The Oncogenic microRNA-27a Targets
978 Genes That Regulate Specificity Protein Transcription Factors and the G2-M Checkpoint in
979 MDA-MB-231 Breast Cancer Cells. *Cancer Research* **67**, 11001-11011 (2007).

980
981 33. Lerner M*, et al.* MiRNA-27a controls FBW7/hCDC4-dependent cyclin E degradation and cell
982 cycle progression. *Cell Cycle* **10**, 2172-2183 (2011).

983
984 34. Ye P*, et al.* Up-regulated MiR-27-3p promotes the G1-S phase transition by targeting
985 inhibitor of growth family member 5 in osteosarcoma. *Biomedicine & Pharmacotherapy* **101**,
986 219-227 (2018).

987
988 35. Su C, Huang DP, Liu JW, Liu WY, Cao YO. miR27a3p regulates proliferation and apoptosis
989 of colon cancer cells by potentially targeting BTG1. *Oncol Lett* **18**, 2825-2834 (2019).

990
991 36. Shang Y*, et al.* The miR27b-CCNG1-P53-miR-508-5p axis regulates multidrug resistance of
992 gastric cancer. *Oncotarget* **7**, (2015).

993
994 37. Lin T, Ma Q, Zhang Y, Zhang H, Yan J, Gao C. MicroRNA-27a functions as an oncogene in
995 human osteosarcoma by targeting CCNG1. *Oncol Lett* **15**, 1067-1071 (2018).

996
997 38. Gibbs JD, Ornoff DM, Igo HA, Zeng JY, Imani F. Cell Cycle Arrest by Transforming Growth
998 Factor β1 Enhances Replication of Respiratory Syncytial Virus in Lung Epithelial Cells. *Journal*
999 *of Virology* **83**, 12424-12431 (2009).

1000
1001 39. Bian T, Gibbs JD, Orvell C, Imani F. Respiratory syncytial virus matrix protein induces lung
1002 epithelial cell cycle arrest through a p53 dependent pathway. *PLoS One* **7**, e38052 (2012).

1003
1004 40. Altan-Bonnet G, Mukherjee R. Cytokine-mediated communication: a quantitative appraisal
1005 of immune complexity. *Nature Reviews Immunology* **19**, 205-217 (2019).

1006
1007 41. Benhamed M, Herbig U, Ye T, Dejean A, Bischof O. Senescence is an endogenous trigger for
1008 microRNA-directed transcriptional gene silencing in human cells. *Nat Cell Biol* **14**, 266-275
1009 (2012).

1010

1011 42. Matsui M, *et al.* Promoter RNA links transcriptional regulation of inflammatory pathway
1012 genes. *Nucleic Acids Res* **41**, 10086-10109 (2013).

1013
1014 43. Di Mauro V, Crasto S, Colombo FS, Di Pasquale E, Catalucci D. Wnt signalling mediates miR-
1015 133a nuclear re-localization for the transcriptional control of Dnmt3b in cardiac cells. *Sci Rep*
1016 **9**, 9320 (2019).

1017
1018 44. Turunen TA, *et al.* Changes in nuclear and cytoplasmic microRNA distribution in response to
1019 hypoxic stress. *Sci Rep* **9**, 10332 (2019).

1020
1021 45. Laitinen P, *et al.* Nuclear microRNA-466c regulates Vegfa expression in response to hypoxia.
1022 *PLOS ONE* **17**, e0265948 (2022).

1023
1024 46. Lou W, Ding B, Fu P. Pseudogene-Derived lncRNAs and Their miRNA Sponging Mechanism in
1025 Human Cancer. *Frontiers in Cell and Developmental Biology* **8**, (2020).

1026
1027 47. Staszak K, Makałowska I. Cancer, Retrogenes, and Evolution. *Life* **11**, 72 (2021).

1028
1029 48. Gray-Owen SD, Blumberg RS. CEACAM1: contact-dependent control of immunity. *Nature*
1030 *Reviews Immunology* **6**, 433-446 (2006).

1031
1032 49. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in
1033 mammalian mRNAs. *Elife* **4**, (2015).

1034
1035 50. Santhakumar D, *et al.* Combined agonist-antagonist genome-wide functional screening
1036 identifies broadly active antiviral microRNAs. *Proc Natl Acad Sci U S A* **107**, 13830-13835
1037 (2010).

1038
1039 51. McCaskill JL, *et al.* Broad-Spectrum Inhibition of Respiratory Virus Infection by MicroRNA
1040 Mimics Targeting p38 MAPK Signaling. *Mol Ther Nucleic Acids* **7**, 256-266 (2017).

1041
1042 52. Pazdrak K, *et al.* MAPK activation is involved in posttranscriptional regulation of RSV-induced
1043 RANTES gene expression. *American Journal of Physiology-Lung Cellular and Molecular*
1044 *Physiology* **283**, L364-L372 (2002).

1045
1046 53. Thomas KW, Monick MM, Staber JM, Yarovinsky T, Carter AB, Hunninghake GW. Respiratory
1047 Syncytial Virus Inhibits Apoptosis and Induces NF-κB Activity through a Phosphatidylinositol
1048 3-Kinase-dependent Pathway*. *Journal of Biological Chemistry* **277**, 492-501 (2002).

1049
1050 54. Gibbs JD, Ornoff DM, Igo HA, Zeng JY, Imani F. Cell cycle arrest by transforming growth factor
1051 beta1 enhances replication of respiratory syncytial virus in lung epithelial cells. *J Virol* **83**,
1052 12424-12431 (2009).

1053

1054 55. Krzyzaniak MA, Zumstein MT, Gerez JA, Picotti P, Helenius A. Host cell entry of respiratory
1055 syncytial virus involves macropinocytosis followed by proteolytic activation of the F protein.
1056 *PLoS Pathog* **9**, e1003309 (2013).

1057
1058 56. Preugschas HF, Hrincius ER, Mewis C, Tran GVQ, Ludwig S, Ehrhardt C. Late activation of the
1059 Raf/MEK/ERK pathway is required for translocation of the respiratory syncytial virus F
1060 protein to the plasma membrane and efficient viral replication. *Cellular Microbiology* **21**,
1061 e12955 (2019).

1062
1063 57. Chen L-F*, et al.* Inhibitory Effect of PIK-24 on Respiratory Syncytial Virus Entry by Blocking
1064 Phosphatidylinositol-3 Kinase Signaling. *Antimicrobial Agents and Chemotherapy* **64**, e00608-
1065 00620 (2020).

1066
1067 58. Daneshpour H, Youk H. Modeling cell–cell communication for immune systems across space
1068 and time. *Current Opinion in Systems Biology* **18**, 44-52 (2019).

1069
1070 59. Stavast CJ, Erkeland SJ. The Non-Canonical Aspects of MicroRNAs: Many Roads to Gene
1071 Regulation. *Cells* **8**, (2019).

1072
1073 60. Zhao F*, et al.* MicroRNA-26b Inhibits Hepatitis B Virus Transcription and Replication by
1074 Targeting the Host Factor CHORDC1 Protein*. *Journal of Biological Chemistry* **289**, 35029-
1075 35041 (2014).

1076
1077 61. Jia X, Bi Y, Li J, Xie Q, Yang H, Liu W. Cellular microRNA miR-26a suppresses replication of
1078 porcine reproductive and respiratory syndrome virus by activating innate antiviral immunity.
1079 *Sci Rep* **5**, 10651 (2015).

1080
1081 62. Gao S, Li J, Song L, Wu J, Huang W. Influenza A virus-induced downregulation of miR-26a
1082 contributes to reduced IFNalpha/beta production. *Virol Sin* **32**, 261-270 (2017).

1083
1084 63. Zhang J, Li Z, Huang J, Yin H, Tian J, Qu L. miR-26a Inhibits Feline Herpesvirus 1 Replication by
1085 Targeting SOCS5 and Promoting Type I Interferon Signaling. *Viruses* **12**, 2 (2020).

1086
1087 64. Liu S, Gao L, Wang X, Xing Y. Respiratory syncytial virus infection inhibits TLR4 signaling via
1088 up-regulation of miR-26b. *Cell Biol Int* **39**, 1376-1383 (2015).

1089
1090 65. Hallak LK, Spillmann D, Collins PL, Peeples ME. Glycosaminoglycan Sulfation Requirements
1091 for Respiratory Syncytial Virus Infection. *Journal of Virology* **74**, 10508 (2000).

1092
1093 66. Guerrero-Plata A*, et al.* Differential Response of Dendritic Cells to Human Metapneumovirus
1094 and Respiratory Syncytial Virus. *American Journal of Respiratory Cell and Molecular Biology*
1095 **34**, 320-329 (2006).

1096

1097 67. Kwok S*, et al.* Effects of primer-template mismatches on the polymerase chain reaction:
1098 Human immunodeficiency virus type 1 model studies. *Nucleic Acids Research* **18**, 999-1005
1099 (1990).

1100
1101 68. Bru D, Martin-Laurent F, Philippot L. Quantification of the Detrimental Effect of a Single
1102 Primer-Template Mismatch by Real-Time PCR Using the 16S rRNA Gene as an Example.
1103 *Applied and Environmental Microbiology* **74**, 1660-1663 (2008).

1104
1105 69. Stadhouders R, Pas SD, Anber J, Voermans J, Mes THM, Schutten M. The Effect of Primer-
1106 Template Mismatches on the Detection and Quantification of Nucleic Acids Using the 5
1107 Nuclease Assay. *The Journal of Molecular Diagnostics* **12**, 109-117 (2010).

1108
1109 70. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time
1110 quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).

1111
1112 71. Burke JM, Sullivan CS. DUSP11 – An RNA phosphatase that regulates host and viral non-
1113 coding RNAs in mammalian cells. *RNA Biology* **14**, 1457-1465 (2017).

1114
1115 72. Gagnon Keith T, Li L, Chu Y, Janowski Bethany A, Corey David R. RNAi Factors Are Present
1116 and Active in Human Cell Nuclei. *Cell Reports* **6**, 211-221 (2014).

1117
1118 73. Schindelin J*, et al.* Fiji: an open-source platform for biological-image analysis. *Nature*
1119 *Methods* **9**, 676-682 (2012).

1120
1121 74. Stael S, Miller LP, Fernández-Fernández ÁD, Van Breusegem F. Detection of Damage-
1122 Activated Metacaspase ActivityActivitiesby Western Blot in Plants. In: *Plant Proteases and*
1123 *Plant Cell Death: Methods and Protocols* (ed^(eds Klemenčič M, Stael S, Huesgen PF).
1124 Springer US (2022).

1125
1126 75. Pall GS, Hamilton AJ. Improved northern blot method for enhanced detection of small RNA.
1127 *Nat Protoc* **3**, 1077-1084 (2008).

1128
1129 76. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1130 *2011* **17**, 3 (2011).

1131
1132 77. Zhao S*, et al.* QuickMIRSeq: a pipeline for quick and accurate quantification of both known
1133 miRNAs and isomiRs by jointly processing multiple samples from microRNA sequencing. *BMC*
1134 *Bioinformatics* **18**, 180 (2017).

1135
1136 78. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**,
1137 357-359 (2012).

1138
1139 79. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features.
1140 *Bioinformatics* **26**, 841-842 (2010).

1141
1142 80. Mann M, Wright PR, Backofen R. IntaRNA 2.0: enhanced and customizable prediction of
1143 RNA–RNA interactions. *Nucleic Acids Research* **45**, W435-W439 (2017).

1144
1145 81. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
1146 *Bioinformatics* **30**, 2114-2120 (2014).

1147
1148 82. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and
1149 genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, 907-915 (2019).

1150
1151 83. Lawrence M*, et al.* Software for Computing and Annotating Genomic Ranges. *PLOS*
1152 *Computational Biology* **9**, e1003118 (2013).

1153
1154 84. Lawrence M, Gentleman R, Carey V. rtracklayer: an R package for interfacing with genome
1155 browsers. *Bioinformatics* **25**, 1841-1842 (2009).

1156
1157 85. Robinson JT*, et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).

1158
1159 86. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for
1160 alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* **47**, e47-e47
1161 (2019).

1162

1163

Figure 1

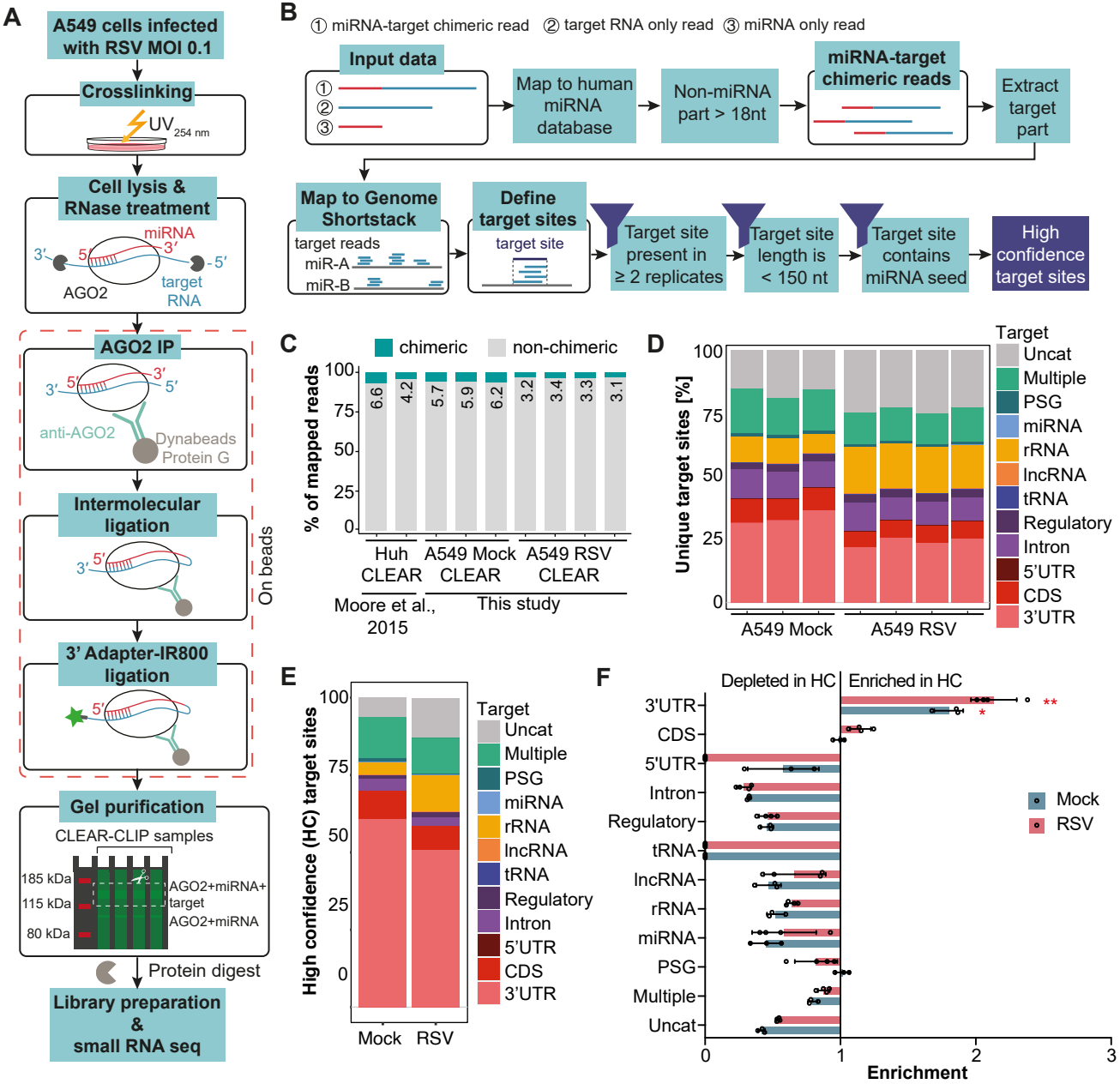

A
A549 cells infected with RSV MOI 0.1
Crosslinking
UV 254 nm
Cell lysis & RNase treatment
miRNA
3'
5'
AGO2
target RNA
AGO2 IP
anti-AGO2
Dynabeads Protein G
Intermolecular ligation
On beads
3' Adapter-IR800 ligation
Gel purification
CLEAR-CLIP samples
185 kDa
115 kDa
80 kDa
AGO2+miRNA+ target
AGO2+miRNA
Protein digest
Library preparation & small RNA seq
B
① miRNA-target chimeric read ② target RNA only read ③ miRNA only read
Input data
Map to human miRNA database
Non-miRNA part > 18nt
miRNA-target chimeric reads
Extract target part
Map to Genome Shortstack
target reads
miR-A
miR-B
Define target sites
target site
Target site present in ≥ 2 replicates
Target site length is < 150 nt
Target site contains miRNA seed
High confidence target sites
C
chimeric
non-chimeric
% of mapped reads
6.6
4.2
5.7
5.9
6.2
3.2
3.4
3.3
3.1
Huh CLEAR
A549 Mock CLEAR
A549 RSV CLEAR
Moore et al., 2015
This study
D
Unique target sites [%]
A549 Mock
A549 RSV
Target
Uncat
Multiple
PSG
miRNA
rRNA
lncRNA
tRNA
Regulatory
Intron
5'UTR
CDS
3'UTR
E
High confidence (HC) target sites
Mock
RSV
F
Depleted in HC
Enriched in HC
Mock
RSV
Enrichment

Figure 2

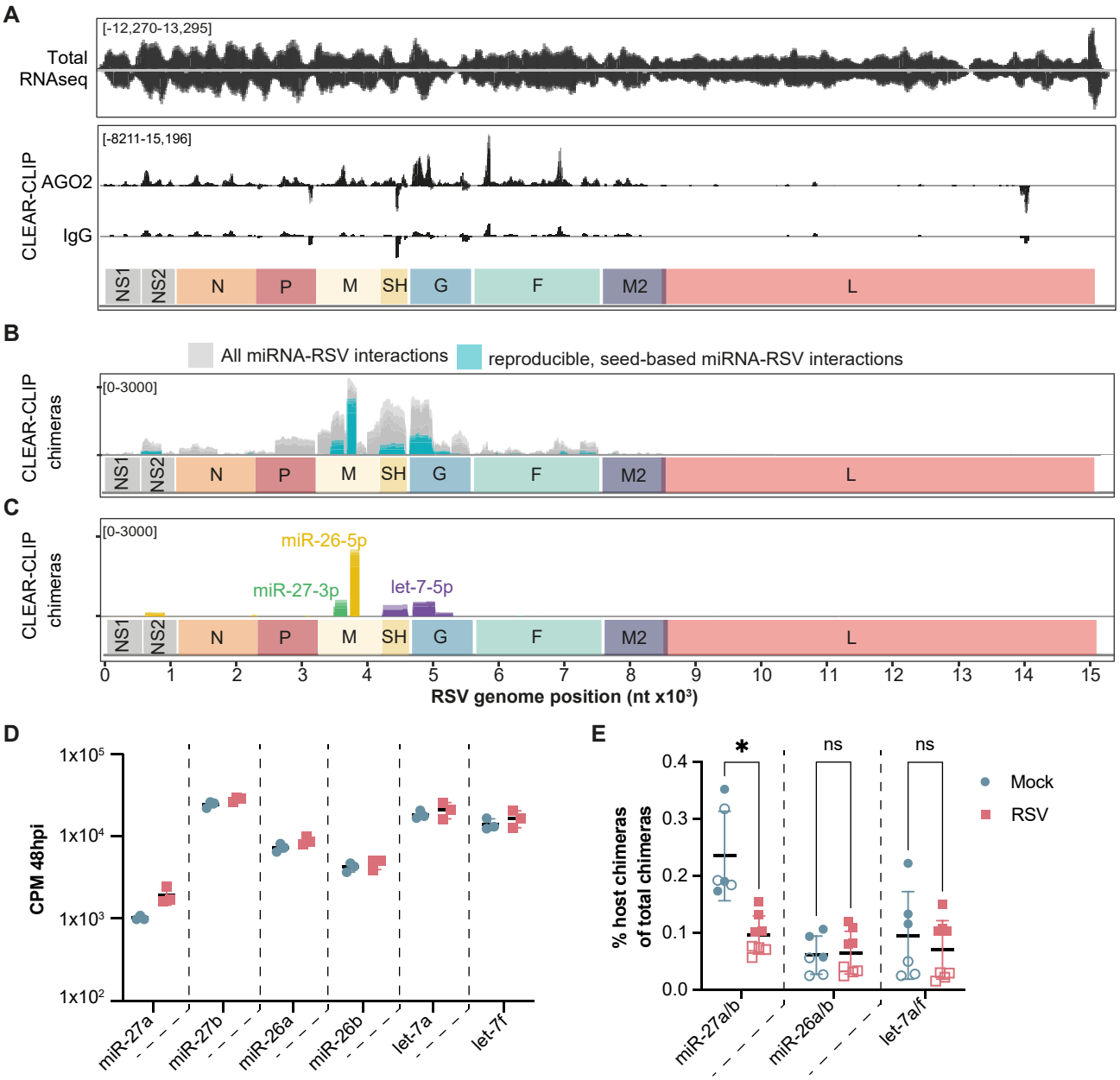

A
[-12,270-13,295]
Total RNAseq
[-8211-15,196]
CLEAR-CLIP
AGO2
IgG
NS1
NS2
N
P
M
SH
G
F
M2
L
B
All miRNA-RSV interactions
reproducible, seed-based miRNA-RSV interactions
CLEAR-CLIP chimeras
[0-3000]
C
miR-26-5p
miR-27-3p
let-7-5p
RSV genome position (nt x10³)
D
CPM 48hpi
1x10⁵
1x10⁴
1x10³
1x10²
miR-27a
miR-27b
miR-26a
miR-26b
let-7a
let-7f
E
% host chimeras of total chimeras
0.4
0.3
0.2
0.1
0.0
*
ns
ns
miR-27a/b
miR-26a/b
let-7a/f
Mock
RSV

Figure 3

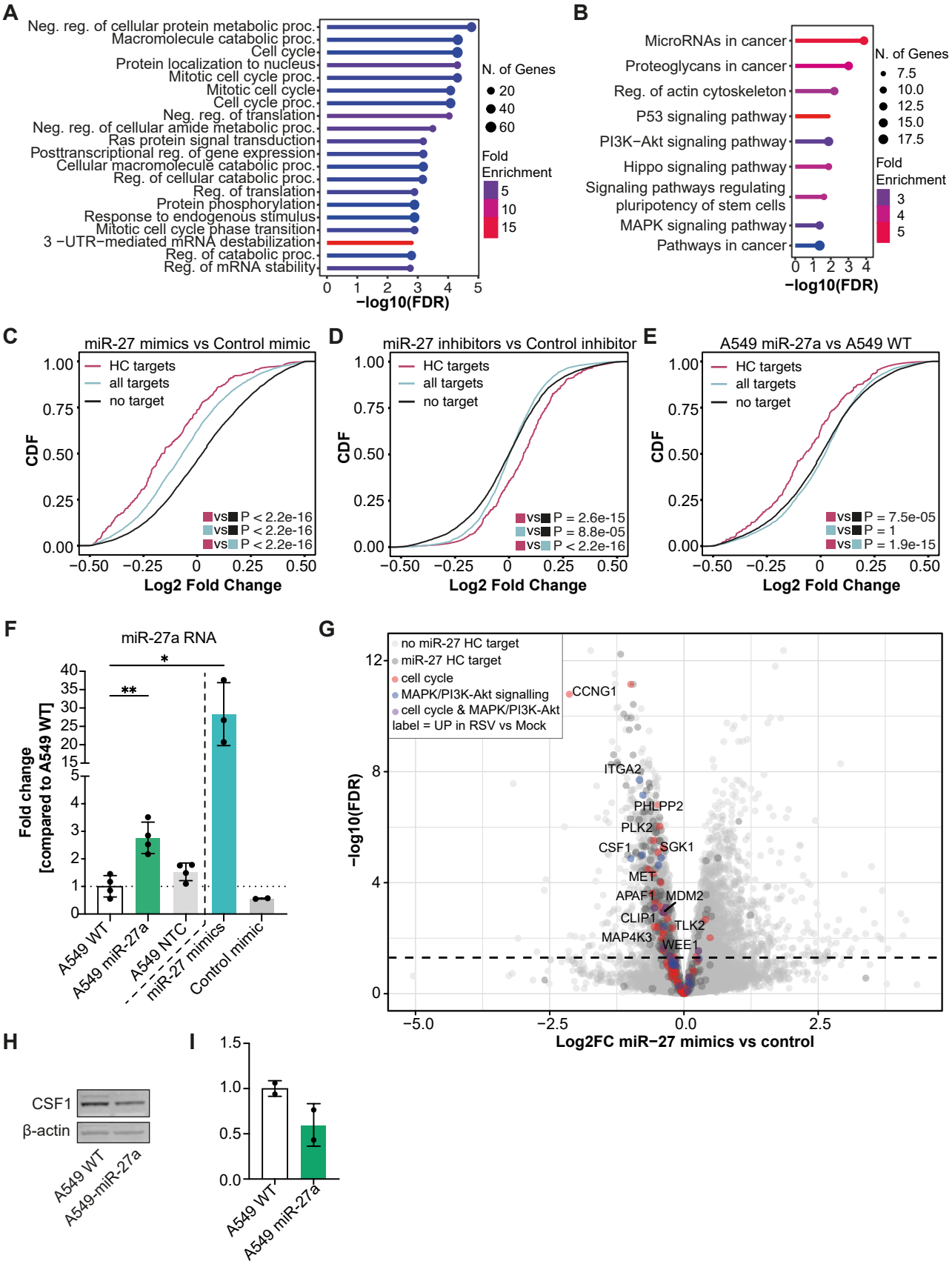

A
Neg. reg. of cellular protein metabolic proc.
Macromolecule catabolic proc.
Cell cycle
Protein localization to nucleus
Mitotic cell cycle proc.
Mitotic cell cycle
Cell cycle proc.
Neg. reg. of translation
Neg. reg. of cellular amide metabolic proc.
Ras protein signal transduction
Posttranscriptional reg. of gene expression
Cellular macromolecule catabolic proc.
Reg. of cellular catabolic proc.
Reg. of translation
Protein phosphorylation
Response to endogenous stimulus
Mitotic cell cycle phase transition
3 −UTR−mediated mRNA destabilization
Reg. of catabolic proc.
Reg. of mRNA stability
0 1 2 3 4 5
−log10(FDR)
N. of Genes
20
40
60
Fold Enrichment
5
10
15
B
MicroRNAs in cancer
Proteoglycans in cancer
Reg. of actin cytoskeleton
P53 signaling pathway
PI3K−Akt signaling pathway
Hippo signaling pathway
Signaling pathways regulating pluripotency of stem cells
MAPK signaling pathway
Pathways in cancer
0 1 2 3 4
−log10(FDR)
N. of Genes
7.5
10.0
12.5
15.0
17.5
Fold Enrichment
3
4
5
C
miR-27 mimics vs Control mimic
HC targets
all targets
no target
CDF
1.00
0.75
0.50
0.25
0.00
vs P < 2.2e-16
vs P < 2.2e-16
vs P < 2.2e-16
−0.50 −0.25 0 0.25 0.50
Log2 Fold Change
D
miR-27 inhibitors vs Control inhibitor
HC targets
all targets
no target
CDF
vs P = 2.6e-15
vs P = 8.8e-05
vs P < 2.2e-16
Log2 Fold Change
E
A549 miR-27a vs A549 WT
HC targets
all targets
no target
CDF
vs P = 7.5e-05
vs P = 1
vs P = 1.9e-15
Log2 Fold Change
F
miR-27a RNA
*
**
Fold change [compared to A549 WT]
40
35
30
25
20
15
5
4
3
2
1
0
A549 WT
A549 miR-27a
A549 NTC
miR-27 mimics
Control mimic
G
no miR-27 HC target
miR-27 HC target
cell cycle
MAPK/PI3K-Akt signalling
cell cycle & MAPK/PI3K-Akt
label = UP in RSV vs Mock
CCNG1
ITGA2
PHLPP2
PLK2
CSF1
SGK1
MET
APAF1
MDM2
CLIP1
TLK2
MAP4K3
WEE1
−log10(FDR)
12
8
4
0
−5.0 −2.5 0.0 2.5
Log2FC miR−27 mimics vs control
H
CSF1
β-actin
A549 WT
A549-miR-27a
I
1.5
1.0
0.5
0.0
A549 WT
A549 miR-27a

Figure 4

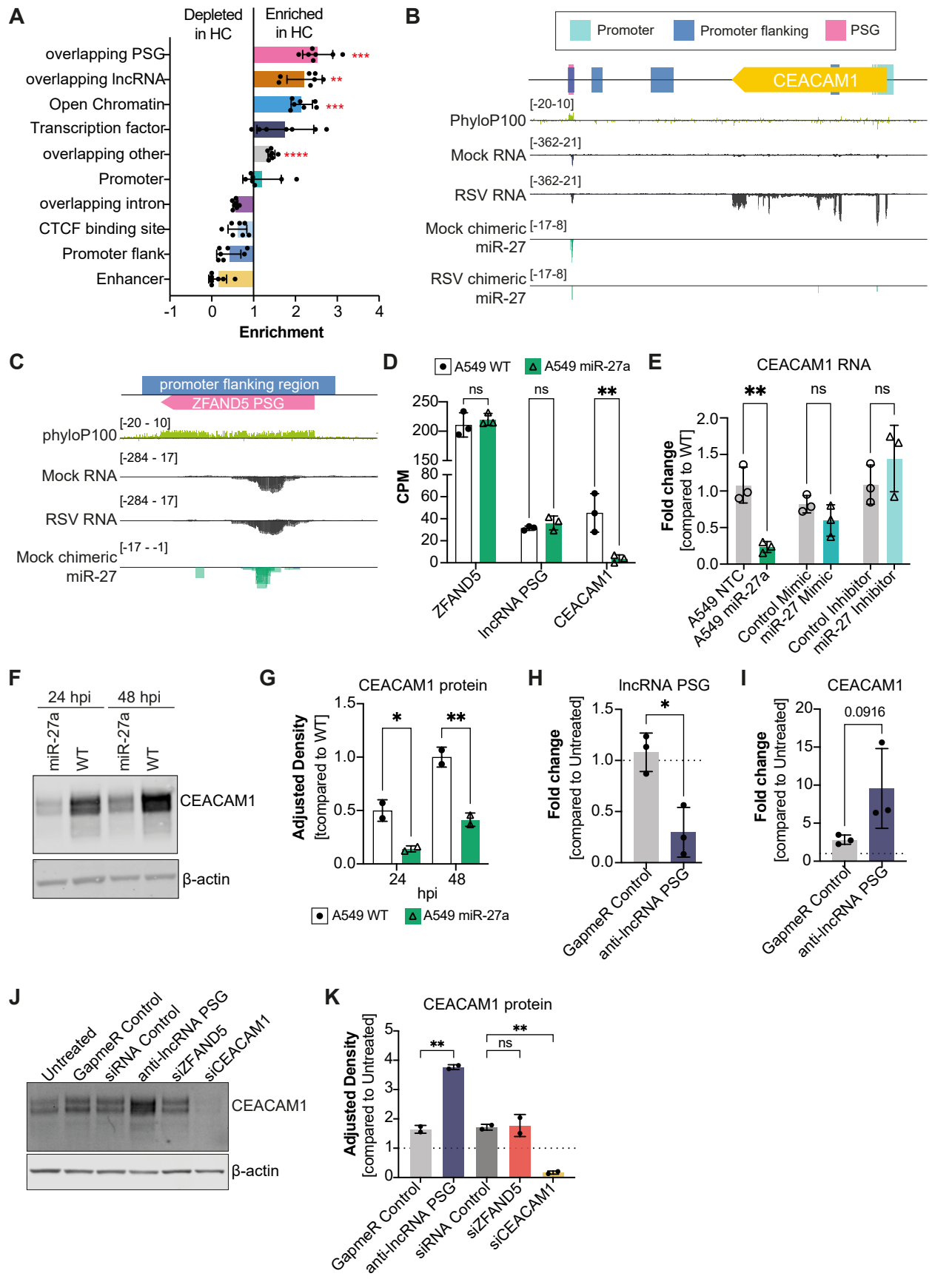
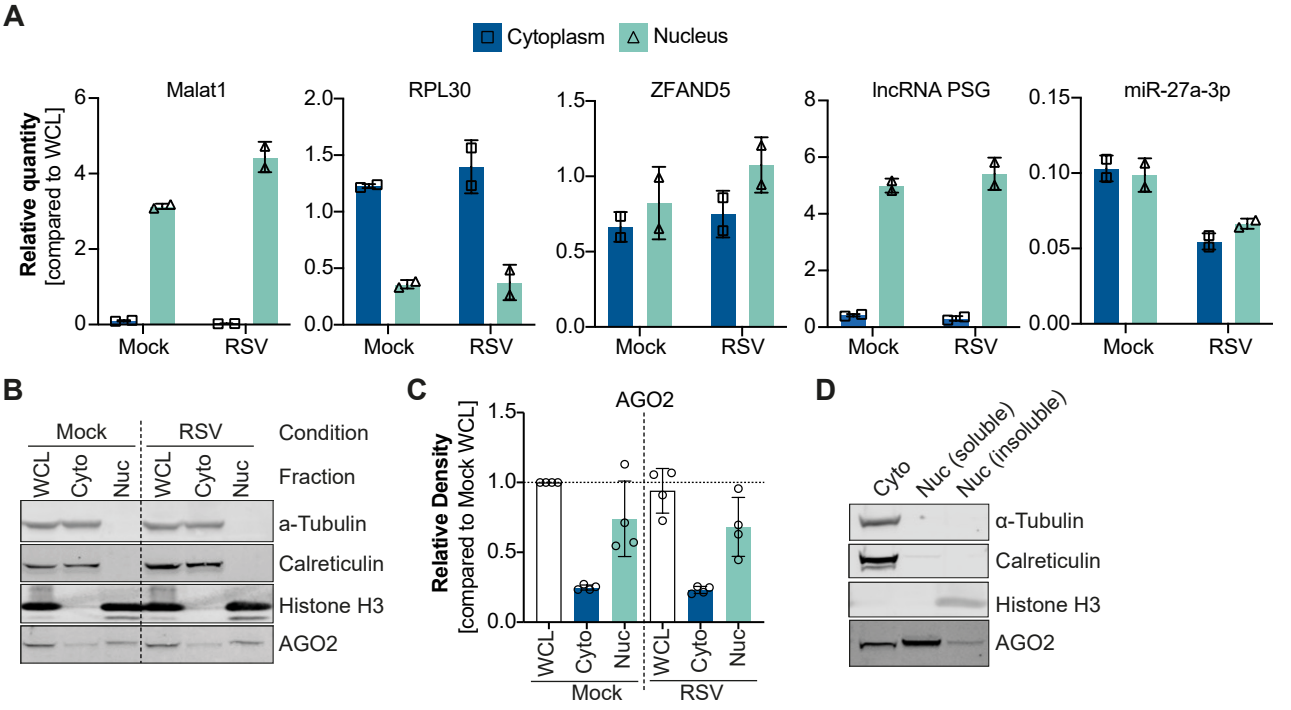

A
Depleted in HC
Enriched in HC
overlapping PSG
overlapping lncRNA
Open Chromatin
Transcription factor
overlapping other
Promoter
overlapping intron
CTCF binding site
Promoter flank
Enhancer
Enrichment
B
Promoter
Promoter flanking
PSG
CEACAM1
PhyloP100
Mock RNA
RSV RNA
Mock chimeric miR-27
RSV chimeric miR-27
C
promoter flanking region
ZFAND5 PSG
phyloP100
Mock RNA
RSV RNA
Mock chimeric miR-27
D
A549 WT
A549 miR-27a
CPM
ZFAND5
lncRNA PSG
CEACAM1
E
CEACAM1 RNA
Fold change [compared to WT]
A549 NTC
A549 miR-27a
Control Mimic
miR-27 Mimic
Control Inhibitor
miR-27 Inhibitor
F
24 hpi
48 hpi
miR-27a
WT
CEACAM1
β-actin
G
CEACAM1 protein
Adjusted Density [compared to WT]
hpi
H
lncRNA PSG
Fold change [compared to Untreated]
GapmeR Control
anti-lncRNA PSG
I
CEACAM1
0.0916
J
Untreated
GapmeR Control
siRNA Control
anti-lncRNA PSG
siZFAND5
siCEACAM1
K
CEACAM1 protein
Adjusted Density [compared to Untreated]

Figure 5


A
Cytoplasm
Nucleus
Relative quantity
[compared to WCL]
Malat1
RPL30
ZFAND5
lncRNA PSG
miR-27a-3p
Mock
RSV
B
Mock
RSV
Condition
WCL
Cyto
Nuc
Fraction
a-Tubulin
Calreticulin
Histone H3
AGO2
C
AGO2
Relative Density
[compared to Mock WCL]
D
Cyto
Nuc (soluble)
Nuc (insoluble)
α-Tubulin
Calreticulin
Histone H3
AGO2