Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Single-cell viral capture identifies injury-remodeling programs and host-response candidates in RSV-infected airway organoids

Rongrong Liu
School of Basic Medical Sciences, Fourth Military Medical University
Ziyu Liu
School of Basic Medical Sciences, Fourth Military Medical University
Pengbo Wang
Xijing Hospital, Fourth Military Medical University
Yi Zhang
Xijing Hospital, Fourth Military Medical University
Shiyuan Hou
School of Basic Medical Sciences, Fourth Military Medical University
Chenyu Zhang
Xijing Hospital, Fourth Military Medical University
Weina Guan
Xijing Hospital, Fourth Military Medical University
Shuhan Wang
School of Basic Medical Sciences, Fourth Military Medical University
Xing Sun
School of Basic Medical Sciences, Fourth Military Medical University
Chenxi Zhao
Xijing Hospital, Fourth Military Medical University
Jinli Huang
Xijing Hospital, Fourth Military Medical University
Wei Zhang
Xijing Hospital, Fourth Military Medical University
Qiuhong Li
Xijing Hospital, Fourth Military Medical University
Yao Zhang
Xijing Hospital, Fourth Military Medical University
Huajie Wu
Xijing Hospital, Fourth Military Medical University
Ruixue Ma

Xijing Hospital, Fourth Military Medical University

Xing-an Wu

School of Basic Medical Sciences, Fourth Military Medical University

Xin Sun

sunxin6@fmmu.edu.cn

Xijing Hospital, Fourth Military Medical University

Research Article

Keywords: respiratory syncytial virus, airway organoids, single-cell RNA sequencing, viral transcript capture, host-response prioritization, adenosine metabolism, SYK, drug-signature reversal

Posted Date: July 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10121104/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Respiratory syncytial virus (RSV) causes substantial pediatric lower respiratory tract disease, but the host programs that separate directly infected airway cells from exposed bystander cells are still not fully resolved. We reanalyzed public single-cell RNA sequencing data from RSV-infected human induced pluripotent stem cell-derived respiratory organoids and used viral transcript capture to distinguish mock cells, RSV RNA-negative bystander cells and RSV RNA-positive cells. RSV-positive cells made up 12.61% of the atlas and showed viral-load-linked stress, epithelial injury-repair, paracrine remodeling, CD73/adenosine and metabolic programs, not merely stronger interferon responses. Bystander cells, by contrast, were enriched for interferon-stimulated and antiviral sensing states. Curated ligand-receptor analysis, exploratory virtual perturbation, multi-layer host-target ranking and drug-signature reversal highlighted IL11, PSAT1, GREM1, PTGS2, SERPINE1, NT5E/CD73 and TNFRSF12A/Fn14 as candidate host nodes. External organoid bulk RNA-seq (GSE263272), antibody-treated organoid scRNA-seq used as a treatment-endpoint reference (GSE263271), primary airway epithelial data (GSE32139) and an in-house RSV time-course qPCR screen supported the injury-remodeling model, while showing that individual candidates vary by context. In silico perturbation placed NT5E/CD73 among network-relevant secondary-tier nodes and ranked the SYK inhibitor Fostamatinib/R406 highly in drug-signature reversal. These findings set priorities for experimental follow-up but do not establish functional rescue or therapeutic efficacy. Together, the study defines a focused set of RSV organoid host-response candidates for functional screening.

Introduction

Respiratory syncytial virus (RSV) remains a major cause of acute lower respiratory tract disease in infants, young children and vulnerable adults. Severe disease reflects not only viral replication but also epithelial injury, innate immune activation, mucus and barrier dysfunction, and tissue remodeling. Human airway organoids and differentiated airway epithelial cultures provide experimentally tractable systems that retain key features of the respiratory epithelium and are increasingly used to study RSV pathogenesis and evaluate antiviral or host-directed interventions.

Bulk transcriptomic comparisons of RSV-infected and mock-treated cultures have identified interferon-stimulated genes, inflammatory mediators and stress pathways. However, bulk profiles average directly RSV transcript-positive cells with exposed RSV transcript-undetected bystander cells. This mixture can obscure whether a candidate host program tracks with intracellular viral burden, reflects bystander antiviral signaling, or represents a change in cell-state composition. Single-cell RNA sequencing with capture of viral transcripts offers a direct way to separate these components within the same infected culture.

We therefore reanalyzed public scRNA-seq data from RSV-infected human induced pluripotent stem cell-derived respiratory organoids (GSE260752). Cells were stratified by viral transcript capture into mock, RSV RNA-negative bystander and RSV RNA-positive populations, then analyzed by cell-state annotation,

pathway scoring, viral-load gradients, curated ligand-receptor prioritization, virtual perturbation and drug-signature reversal. To ask whether the host programs were specific to the discovery dataset, we incorporated external organoid bulk RNA-seq (GSE263272), antibody-treated organoid scRNA-seq (GSE263271), a primary polarized airway epithelial dataset (GSE32139) and an in-house RSV qPCR time course. The goal was to identify host-response mechanisms that remain consistent across viral burden, cell state, external datasets and computational prioritization, while leaving interventional claims to future wet-lab testing.

Results

Single-cell viral transcript capture separates infected and bystander cell states

Quality-controlled GSE260752 cells formed a respiratory organoid atlas containing 17,496 cells: 8,731 mock cells, 6,559 RSV-exposed cells without detectable RSV RNA and 2,206 RSV RNA-positive cells (Fig. 1). Using RSV transcript counts as a direct viral-capture feature, 49.90% of cells were assigned to the mock state, 37.49% to the RSV-bystander state and 12.61% to the RSV-positive state. This classification allowed infected cells and bystanders to be compared within the same organoid infection experiment, reducing the ambiguity inherent to bulk RSV versus mock contrasts. Quality-control metrics and the RSV-capture boundary further supported the infection-state definition used throughout the analysis (Supplementary Fig. S7).

Single-cell viral transcript capture is expected to be specific when RSV UMIs are detected, but it can miss low-abundance viral RNA. We therefore used an operational definition: bystander cells were RSV-exposed cells without detectable viral transcripts, not cells proven to be completely uninfected. This definition preserves a high-specificity RSV-positive compartment while allowing for missed low-viral-load infected cells.

We also evaluated stricter RSV UMI thresholds as a sensitivity analysis (Supplementary Data 2). Among RSV-exposed cells, thresholds of RSV UMI > = 2, >=3, >=5 and > = 10 retained 826, 569, 461 and 397 RSV transcript-positive cells, respectively. Across these stricter thresholds, stress-remodeling, epithelial injury-repair and metabolic-rewiring module differences between threshold-positive and threshold-negative exposed cells remained directionally stronger than at the UMI > = 1 threshold, whereas IFN/ISG scores did not show the same monotonic increase. This supports the interpretation that the infected-cell injury-remodeling signal is not driven solely by single-UMI assignments.

Refined annotation identifies RSV-enriched epithelial injury states

Subclustering and marker-based annotation resolved epithelial and non-epithelial compartments and further identified IFN-high and stress-injury cell states (Fig. 2 and Supplementary Fig. S1). RSV RNA-positive cells were not uniformly distributed across the atlas; instead, viral signal was enriched in IFN-

high and stress-injury states. Within epithelial subsets, RSV burden was concentrated in a limited set of susceptible states, consistent with viral tropism and infection-induced transcriptional remodeling (Supplementary Fig. S2). The numerically labeled epithelial subclusters with the highest RSV-positive fractions were subclusters 8, 12 and 6. Marker-score annotation mapped these clusters mainly to IFN-high/stress-injury, AT2-like/cycling epithelial and pan-epithelial/IFN-high/stress-injury mixed states, respectively, providing a biological interpretation for the subcluster labels shown in the supplementary figures.

Directly infected cells and bystanders show divergent host-response programs

Module scoring separated two major response modes. RSV-bystander cells showed prominent interferon-stimulated and pattern-recognition receptor programs, including canonical antiviral sensing features. By contrast, RSV-positive cells showed stronger stress-remodeling, epithelial injury-repair, paracrine growth-factor and metabolic-rewiring signatures (Fig. 3). In RSV-positive versus bystander cells, the mean differences for infected-cell stress remodeling, epithelial injury repair and metabolic rewiring were 0.079, 0.107 and 0.049 module-score units, respectively, with FDR-adjusted cell-level P values ranging from 1.03e-43 to 6.76e-16. Dose-response ranking further identified genes that increased with RSV RNA burden, linking the infected-cell state to injury-remodeling rather than to interferon signaling alone (Supplementary Fig. S3).

A viral-load gradient links RSV burden to host injury-remodeling

RSV-positive cells were stratified into viral-load groups according to RSV transcript counts; the high-load group used an RSV UMI cutoff of > = 3 and contained 569 cells (Supplementary Data 2). Across this gradient, stress-remodeling, epithelial injury-repair and metabolic-rewiring scores increased with viral burden (Fig. 4). Spearman analyses across single cells linked RSV UMI to infected-cell stress remodeling (rho = 0.173, FDR = 6.56e-59), epithelial injury repair (rho = 0.171, FDR = 5.62e-58) and metabolic rewiring (rho = 0.103, FDR = 1.08e-21), whereas IFN/ISG association was much smaller (rho = 0.042, FDR = 1.15e-4). Transcription-factor and pathway summaries implicated AP-1/stress response, interferon regulatory programs, NF-kappaB-related inflammatory signaling, TGF-beta/EGFR/TWEAK-associated remodeling and metabolic adaptation (Supplementary Fig. S4). These results support a model in which direct viral burden is coupled to epithelial injury and repair programs that only partially overlap with bystander antiviral responses.

Paracrine remodeling axes nominate interpretable host targets

Curated expression-based ligand-receptor and pathway integration prioritized TGFB/TGFBR2, HBEGF/EGFR, VEGFA/FLT1, IL11/IL6ST and TWEAK/TNFRSF12A axes (Supplementary Fig. S5). These

ligand-receptor results were used as mechanistic context rather than as formal probabilistic cell-cell communication inference, and the scoring inputs are provided in Supplementary Data 2. An integrated target dashboard combined differential expression, viral-load association, module membership and external interpretability (Fig. 5 and Supplementary Fig. S6). The top candidates included IL11, SERPINE1, NT5E/CD73, TNFRSF12A/Fn14, PTGS2, GDF15, GREM1, FOSL1, SPHK1 and serine-biosynthesis genes such as PSAT1. These genes map to injury repair, profibrotic remodeling, adenosine metabolism, TWEAK signaling, prostaglandin biology and metabolic adaptation.

Exploratory computational perturbation prioritizes candidate host-response nodes

As exploratory analyses, we used virtual knockout scoring and drug-signature reversal to rank host nodes whose removal or pharmacologic reversal was computationally predicted to attenuate RSV-associated injury-remodeling programs (Fig. 6 and Supplementary Fig. S8). IL11, PSAT1, PTGS2 and GREM1 were assigned Tier 1 priority because they showed strong predicted network impact and mechanistic links to infected-cell remodeling. GDF15, NT5E/CD73, TNFRSF12A/Fn14 and SERPINE1 were assigned Tier 2 priority because they showed consistent biological relevance and cross-dataset or qPCR support but somewhat weaker or more context-dependent computational perturbation scores. A subsequent post hoc multi-layer reprioritization that merged evidence strength, virtual perturbation, viral-load association, actionability and tier weights retained IL11, PSAT1, GREM1 and PTGS2 as the most stable high-priority host candidates, with GDF15, NT5E/CD73, TNFRSF12A/Fn14 and SERPINE1 forming a biologically coherent secondary tier (Supplementary Fig. S10). The complete ranking matrix and scoring weights are provided in Supplementary Data 2.

Because CD73 and R406 were selected for near-term experimental follow-up, we further summarized the relevant computational support in a dedicated in silico perturbation panel (Supplementary Fig. S13). NT5E/CD73 ranked as a secondary-tier network node in the clean virtual knockout analysis, with a combined prioritization score of 3.58 and 14 secondary affected genes. In the drug-signature reversal analysis, Fostamatinib/R406 ranked second overall among host-response perturbagens, with a host reversal score of 87.5/100 and coverage of IL11, NT5E, TNFRSF12A, SERPINE1, LIF and FOSL1. These results nominate CD73/adenosine and SYK/R406 axes for follow-up testing, but they is best interpreted as computational triage rather than functional proof.

External datasets support the injury-remodeling framework

GSE263272 bulk RNA-seq provided the main external organoid support dataset because it includes mock, RSV-A and RSV-B conditions. Because GSE263272 and GSE260752 are both organoid-based resources, we treat GSE263272 as cross-dataset support rather than as definitive biological independence from the discovery model. RSV-A infection induced multiple discovery candidates, including IL11, SERPINE1, NT5E, TNFRSF12A, GDF15, GREM1, PTGS2, FOSL1 and SPHK1, and

reproduced interferon, injury-repair, TWEAK-EGFR-TGF, CD73/adenosine and metabolic modules (Fig. 7). The antibody-treated GSE263271 scRNA-seq dataset was interpreted conservatively as a post-treatment endpoint reference because it lacks untreated RSV and mock comparators and showed no usable viral-load gradient. Finally, the primary polarized airway epithelial dataset GSE32139 showed induction of antiviral controls and supported TNFRSF12A, NT5E, SERPINE1 and PTGS2 as concordant epithelial response genes, whereas IL11 was not induced in that context (Supplementary Fig. S9).

An additional in-house RSV time-course qPCR screen provided orthogonal wet-lab support for this framework (Supplementary Fig. S12). RSV signal increased from 24 to 72 h, IFIT1 and ISG15 were induced as antiviral controls, and IL11, PTGS2, NT5E, FOSL1 and GDF15 also increased over time. In the same assay, SERPINE1 and TNFRSF12A showed weaker or more variable changes than the strongest antiviral and remodeling readouts. Complete raw Ct values, 18S-normalized dCt/ddCt calculations, replicate-level summaries, primer information and qPCR processing metadata are provided as Supplementary Data 1 and figure-level source data. Together, these results strengthen the view that RSV infection couples antiviral sensing to a partial epithelial injury-remodeling program while also highlighting context dependence among individual candidate genes.

Discussion

This study shows how direct viral transcript capture can refine RSV organoid scRNA-seq analysis. Instead of treating infected cultures as a single bulk-like condition, we separated mock cells, RSV RNA-negative bystanders and RSV RNA-positive cells. The main biological finding is that RSV-positive cells were not simply the strongest interferon responders. They displayed a viral-load-linked injury-remodeling state involving stress response, epithelial repair, paracrine signaling, adenosine metabolism and metabolic rewiring, whereas bystander cells carried stronger interferon-stimulated and antiviral sensing programs.

The candidate genes nominated here are mechanistically plausible. IL11 is linked to epithelial stress and fibro-inflammatory remodeling; SERPINE1 marks tissue injury, matrix remodeling and fibrinolytic imbalance; NT5E/CD73 generates extracellular adenosine and may shape epithelial barrier and immune tone; TNFRSF12A/Fn14 is a damage-inducible receptor in the TWEAK pathway; and PTGS2 connects infection-induced inflammation to prostaglandin biology. Their convergence across viral-load gradients, paracrine analysis and external datasets suggests that RSV pathogenesis in organoids involves an epithelial injury-remodeling layer that is distinct from canonical antiviral interferon signaling. The strengthened integrated ranking also suggests that PSAT1 and GREM1 deserve more attention than would be apparent from a purely cytokine-centered interpretation of the dataset.

The external support strategy is important for interpretation. GSE263272 provides the most direct external organoid comparison because it contains mock and RSV-infected samples, but it should not be over-interpreted as a fully independent in vivo validation system because it remains an organoid transcriptomic dataset. GSE32139 adds a primary airway epithelial reference and supports several core

candidates, especially TNFRSF12A, NT5E, SERPINE1 and PTGS2. In contrast, GSE263271 should not be used as untreated infection support. Its value is as a boundary reference for antibody-treated organoids, emphasizing that post-neutralization host states should not be confused with persistent viral-replication gradients. The in-house qPCR time course adds initial experimental expression support for IFIT1, ISG15, IL11, PTGS2, NT5E, FOSL1 and GDF15, while also underscoring that not every candidate is uniformly robust across assays. In particular, SERPINE1 and TNFRSF12A were weaker or more context-dependent in this first-pass experiment, so the current wet-lab screen is best interpreted as partial support for the broader injury-remodeling framework rather than as uniform confirmation of every prioritized node.

The virtual knockout and perturbagen analyses should be read as prioritization tools. They can reduce a long differential-expression list to a smaller set of experimentally testable nodes, but they do not establish causal antiviral or tissue-protective effects. The integrated target score was not pre-registered and should therefore be viewed as a post hoc prioritization matrix, with all weights and component scores provided in Supplementary Data 2. This is especially important for NT5E/CD73 and Fostamatinib/R406. The current data support choosing them as follow-up candidates because CD73 connects the infected-cell injury program to extracellular adenosine metabolism and because the SYK inhibitor Fostamatinib/R406 is predicted to reverse multiple RSV-positive injury-remodeling genes. However, protein-level CD73 induction, R406 dose-response, cytotoxicity, viral-load reduction and remodeling-marker rescue remain to be tested experimentally. CD73 and R406 were prioritized for near-term follow-up because they combine biological relevance, available reagents and experimental tractability; they does not establish the only or definitively superior candidates.

We deliberately do not advance a direct RSV F-protein inhibitor combination as a main result in this submission-ready version. A separate structure-guided direct-antiviral scaffold was explored during project development, but final docking, trajectory-level molecular-dynamics and resistance-hotspot tables were not available for a strong structure-based claim. We therefore retained the direct-antiviral module only as an evidence-boundary audit in the supplementary material, documenting why it should remain a future validation context rather than a central result here (Supplementary Fig. S11). This framing keeps the manuscript focused on the host-response discovery that is best supported by the present data.

Several limitations should be considered. First, the discovery dataset contains a limited number of conditions and does not by itself establish time-resolved causality. Second, iPSC-derived respiratory organoids are useful epithelial infection models but do not fully reproduce mature in vivo airway architecture, immune-cell recruitment, vascular context or patient-level heterogeneity; therefore, the candidate axes should not be generalized directly to clinical disease without primary-cell, organoid perturbation and in vivo validation. Third, viral transcript capture is influenced by sequencing depth, viral gene annotation and capture efficiency; therefore, RSV-bystander cells is best interpreted as exposed cells without detected viral transcripts rather than as a pure uninfected compartment. Fourth, single-cell discovery and bulk support datasets are not equivalent and may differ because of cell-state composition and shared organoid-model biology. Fifth, some candidates, such as IL11 and serine-biosynthesis genes,

showed context-dependent behavior across external datasets. Sixth, the in-house qPCR experiment is a first-pass time-course expression screen and does not substitute for protein-level or functional validation. Finally, the CD73/adenosine and SYK/R406 axes remain computationally prioritized and experimentally unproven. The appropriate conclusion is therefore that this study identifies a coherent and externally supported RSV host-response program and provides a prioritized resource for functional screening, not that it proves therapeutic efficacy.

Methods

Study design and datasets

This study combined secondary analysis of public transcriptomic datasets with a targeted in-house RSV time-course qPCR support screen. GSE260752 was used as the single discovery scRNA-seq dataset for RSV-infected and mock human induced pluripotent stem cell-derived respiratory organoids. GSE263272 was used as an external organoid bulk RNA-seq support dataset including mock, RSV-A and RSV-B conditions. Because it is also an organoid transcriptomic dataset, it was interpreted as cross-dataset support rather than as fully independent biological validation. GSE263271 was evaluated as an antibody-treated organoid scRNA-seq reference containing RSV plus nirsevimab and RSV plus palivizumab samples. GSE32139 was used as an external primary polarized airway epithelial microarray support dataset with RSV and time-matched mock conditions.

Single-cell preprocessing and annotation

Raw or processed 10x Genomics count matrices were imported into R and analyzed using a Seurat-style workflow. Cells were filtered according to standard quality-control metrics, including detected gene number, UMI count and mitochondrial transcript proportion. Normalization, variable-feature selection, principal-component analysis, UMAP embedding and graph-based clustering were then performed. Broad compartments and refined epithelial states were annotated by marker-gene inspection using curated airway and lung epithelial marker sets. Marker genes for each refined annotation state, including IFN-high and stress-injury programs, are provided in Supplementary Data 2.

RSV transcript capture and infection-state definition

Viral features were identified from RSV-specific transcript annotations present in the count matrix. Cells with detectable RSV UMI signal were defined as RSV-positive cells. Cells from RSV-exposed samples without detectable RSV transcripts were defined as RSV-bystander cells. Cells from mock samples were defined as mock cells. The primary RSV-positive threshold was set at RSV UMI > = 1 to capture cells with direct evidence of viral transcript detection. This rule was used as a high-specificity infection-state definition, while recognizing that droplet-based scRNA-seq can miss low-abundance viral RNA because of sampling depth and capture dropout. Accordingly, RSV-bystander denotes virus-exposed, RSV transcript-undetected cells, not cells proven to be infection-free. Sensitivity summaries were calculated for RSV UMI thresholds > = 2, >=3, >=5 and > = 10, including retained cell counts and mean module-score

differences between threshold-positive and threshold-negative exposed cells. RSV-positive cells were further stratified by RSV UMI counts to test viral-load dose-response relationships.

Differential expression, module scoring and enrichment analysis

Differential expression analyses were performed across infection states and across viral-load strata. Gene modules representing interferon response, pattern-recognition sensing, epithelial injury repair, stress remodeling, paracrine signaling, CD73/adenosine biology and metabolic rewiring were scored at single-cell level. Module comparisons were summarized as mean and median differences between cell states with FDR-adjusted cell-level P values; these P values are descriptive of single-cell distributions and are not interpreted as independent biological-replicate tests. Pathway and transcription-factor summaries were derived from ranked gene lists and interpreted at the module level rather than as proof for individual genes.

Ligand-receptor, virtual perturbation and drug-signature analyses

Ligand-receptor pairs were prioritized using a curated expression-based strategy rather than CellChat or NicheNet probabilistic inference. Candidate axes were restricted to injury-remodeling and epithelial response pairs, and a paracrine score was calculated as the mean ligand expression in RSV-positive sender cells multiplied by the mean receptor expression in RSV-bystander receiver cells. This score was used only for mechanistic context and did not assign formal interaction P values. Virtual knockout scoring was performed with scTenifoldKnk in RSV-positive cells using 1,000 cells, 2,864 genes, nc_nNet = 10 and nc_nCells = 500 across candidate targets IL11, SERPINE1, NT5E, TNFRSF12A, SPHK1, FOSL1, PSAT1, PHGDH, GDF15, LIF, GREM1 and PTGS2. Clean post-processing removed the KO target itself from affected-gene overlap statistics to reduce self-overlap inflation. Drug-signature reversal analysis compared RSV-associated gene signatures with public LINCS/Enrichr perturbation signatures and ranked candidate perturbagens according to reversal strength, FDR and coverage of prioritized host genes. These analyses were used only for candidate prioritization; computational outputs were not interpreted as direct drug-efficacy evidence or true genetic knockout biology.

Integrated host-target reprioritization and in-house qPCR support screen

Host-target evidence was integrated across differential expression, viral-load association, curated module membership, virtual perturbation summaries, previous tier labels and actionability features to derive an exploratory host priority score for experimental triage. The score was not pre-registered and was calculated as 0.30 x normalized transcriptomic evidence, 0.25 x normalized clean virtual-KO score, 0.20 x normalized positive log2FC, 0.15 x actionability score and 0.10 x tier weight. Drug reversal scores used 0.65 x Enrichr/LINCS combined score, 0.20 x -log10(FDR) and 0.15 x prioritized target coverage. Full scoring matrices and weights are provided in Supplementary Data 2. The in-house qPCR screen

evaluated RSV time-course expression patterns for antiviral controls and selected candidate genes across mock, 24 h, 48 h and 72 h conditions. Raw Ct values from the 10-primer and 8-primer qPCR batches were consolidated at replicate level. Target-gene Ct values were normalized to 18S within each batch and replicate to calculate dCt; the mock mean dCt for each gene and batch was used as the reference for ddCt; relative expression was calculated as 2^-ddCt. Ct values > = 35 or missing values were flagged as low-abundance measurements. Group summaries and exploratory two-sided t tests were calculated on dCt values against the mock group. The qPCR output was used as an additional support layer, with emphasis on directionality and consistency rather than definitive validation of every candidate. Complete raw Ct tables, normalized values, summary statistics, primer sequences and processing metadata are provided in Supplementary Data 1.

External support datasets and evidence-boundary audit

For GSE263272, candidate genes and mechanism modules were compared between mock, RSV-A and RSV-B bulk RNA-seq samples. For GSE263271, because untreated RSV and mock comparators were absent and viral-feature capture did not support a usable viral-load gradient, the dataset was interpreted only as an antibody-treated endpoint reference. For GSE32139, RSV 48 h samples were compared with time-matched mock 48 h samples, focusing on candidate genes and antiviral controls. An optional direct-antiviral structure scaffold explored during project development was retained only as a supplementary evidence-boundary audit, because final docking, molecular-dynamics and resistance-hotspot result tables were not available.

Statistical analysis and visualization

Differential expression and support-dataset analyses were performed using the corresponding R and Python scripts in the project workspace. False-discovery-rate adjustment was used where multiple comparisons were performed. Figures were generated in a consistent white-background publication style. Figure-level source data were exported as Excel workbooks when available, including source data for the in silico perturbation prioritization figure.

Declarations

Ethics statement

This study used only public, de-identified datasets and did not involve new participant recruitment, animal experiments or collection of identifiable private information.

Data availability

The public datasets analyzed in this study are available from the Gene Expression Omnibus under accession numbers GSE260752, GSE263271, GSE263272 and GSE32139. Figure-level source data generated for this manuscript are provided in the accompanying source_data directory. Supplementary

Data 1 contains the in-house qPCR raw Ct table, normalized expression calculations, replicate summaries, primer sequences and qPCR processing metadata. Supplementary Data 2 contains infection-state counts, RSV UMI threshold sensitivity analyses, marker annotation support, ligand-receptor scoring inputs, virtual-KO settings, drug-reversal rankings and host-target scoring weights.

Code availability

Scripts used to generate figures and source-data files are located in the project scripts and manuscript directories and should be uploaded as supplementary code at submission.

Acknowledgements

We thank the research teams who generated and publicly shared RSV organoid and airway epithelial transcriptomic datasets.

Author contributions

RRL, ZYL, PBW, YZ, SYH and XingS conceived and designed the study and curated the data. YZ, PBW, CXZ, SHW, WNG, JLH and HJW wrote the first draft of the manuscript. CYZ, QHL, YaoZ and WZ performed the literature screening, verified the references and prepared the figures. RXM, XAW and XinS supervised the study, critically revised the manuscript for important intellectual content, and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by the Shaanxi Provincial Natural Science Basic Research Program (2024JC-ZDXM-42), the Assistance Program of Xijing Hospital (XJZT26JS052 and XJZT26CX04),Clinical Research Project of Air Force Military Medical University (2025LC2514), Xi'an Science and Technology Plan Project(24YXYJ0145), the Natural Science Foundation of Shaanxi Province (2025JC-YBQN-1178).

References

1. Li Y, Wang X, Blau DM, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. Lancet. 2022;399:2047–2064.

2. Hashimoto R, et al. Human iPS cell-derived respiratory organoids as a model for respiratory syncytial virus infection. Life Sci Alliance. 2025;8:e202402837. doi:10.26508/lsa.202402837.
3. Ioannidis I, McNally B, Willette M, Peeples ME, Chaussabel D, Durbin JE, Ramilo O, Mejias A, Flano E. Plasticity and virus specificity of the airway epithelial cell immune response during respiratory virus infection. J Virol. 2012;86:5422–5436. doi:10.1128/JVI.06757-11.
4. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–3587.e29.
5. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. Cell. 2019;177:1888–1902.e21.
6. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36:411–420. doi:10.1038/nbt.4096.
7. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9:5233.
8. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90-W97.
9. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell. 2017;171:1437–1452.e17.
10. Duan Q, Reid SP, Clark NR, et al. L1000CDS2: LINCS L1000 characteristic direction signatures search engine. NPJ Syst Biol Appl. 2016;2:16015.
11. Osorio D, Zhong Y, Li G, et al. scTenifoldNet and scTenifoldKnk: network-based virtual knockout tools for single-cell gene regulatory analysis. Patterns. 2022;3:100434.
12. Openshaw PJM, Chiu C, Culley FJ, Johansson C. Protective and harmful immunity to RSV infection. Annu Rev Immunol. 2017;35:501–532. doi:10.1146/annurev-immunol-051116-052206.
13. Pickles RJ. Human airway epithelial cell cultures for modeling respiratory syncytial virus infection. Curr Top Microbiol Immunol. 2013;372:371–387.
14. San-Juan-Vergara H, Sampayo-Escobar V, Reyes N, et al. Cholesterol-rich microdomains as docking platforms for respiratory syncytial virus in normal human bronchial epithelial cells. J Virol. 2012;86:1832–1843.
15. Shi T, McAllister DA, O'Brien KL, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet. 2017;390:946–958.
16. Plasschaert LW, Zilionis R, Choo-Wing R, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature. 2018;560:377–381. doi:10.1038/s41586-018-0394-6.
17. Travaglini KJ, Nabhan AN, Penland L, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature. 2020;587:619–625. doi:10.1038/s41586-020-2922-4.

18. Schafer S, Viswanathan S, Widjaja AA, et al. IL-11 is a crucial determinant of cardiovascular fibrosis. Nature. 2017;552:110–115.
19. Ng B, Dong J, D'Agostino G, et al. Interleukin-11 is a therapeutic target in idiopathic pulmonary fibrosis. Sci Transl Med. 2019;11:eaaw1237.
20. Hasko G, Linden J, Cronstein B, Pacher P. Adenosine receptors: therapeutic aspects for inflammatory and immune diseases. Nat Rev Drug Discov. 2008;7:759–770. doi:10.1038/nrd2638.
21. Idzko M, Ferrari D, Eltzschig HK. Nucleotide signalling during inflammation. Nature. 2014;509:310–317. doi:10.1038/nature13085.
22. Braselmann S, Taylor V, Zhao H, et al. R406, an orally available spleen tyrosine kinase inhibitor, blocks Fc receptor signaling and reduces immune complex-mediated inflammation. J Pharmacol Exp Ther. 2006;319:998–1008.

Figures

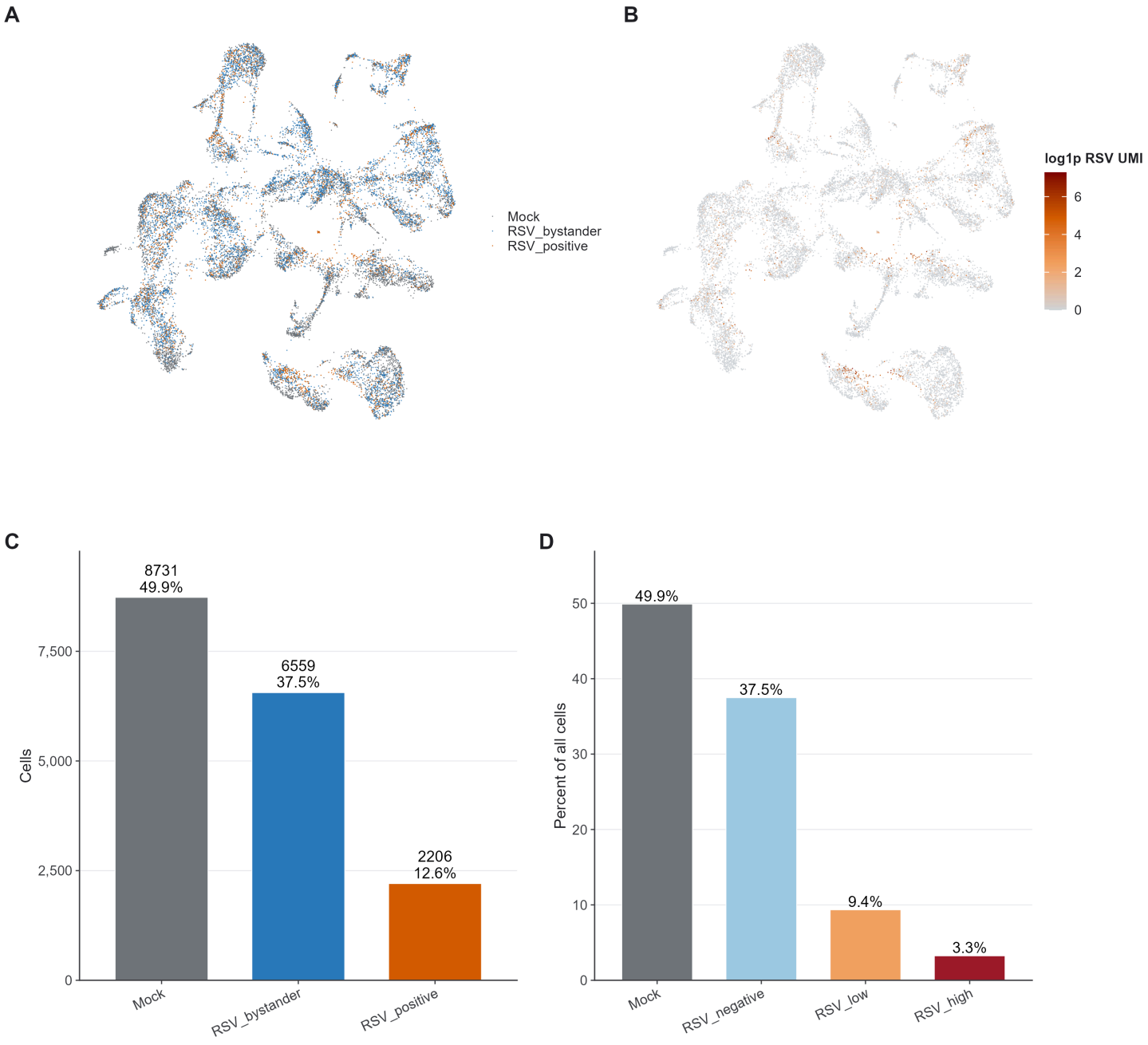


Figure 1

RSV infection atlas and direct viral transcript capture. (A) UMAP projection of the organoid single-cell atlas colored by mock, RSV RNA-negative bystander and RSV RNA-positive infection states. (B) Single-cell RSV transcript signal overlaid on the same embedding. (C) Infection-state cell counts, including 8,731 mock cells, 6,559 RSV transcript-undetected bystander cells and 2,206 RSV transcript-positive cells. (D) RSV UMI-based viral-load strata used for downstream gradient analyses. The figure supports direct separation of RSV transcript-positive cells from exposed transcript-undetected cells within the same infection experiment.

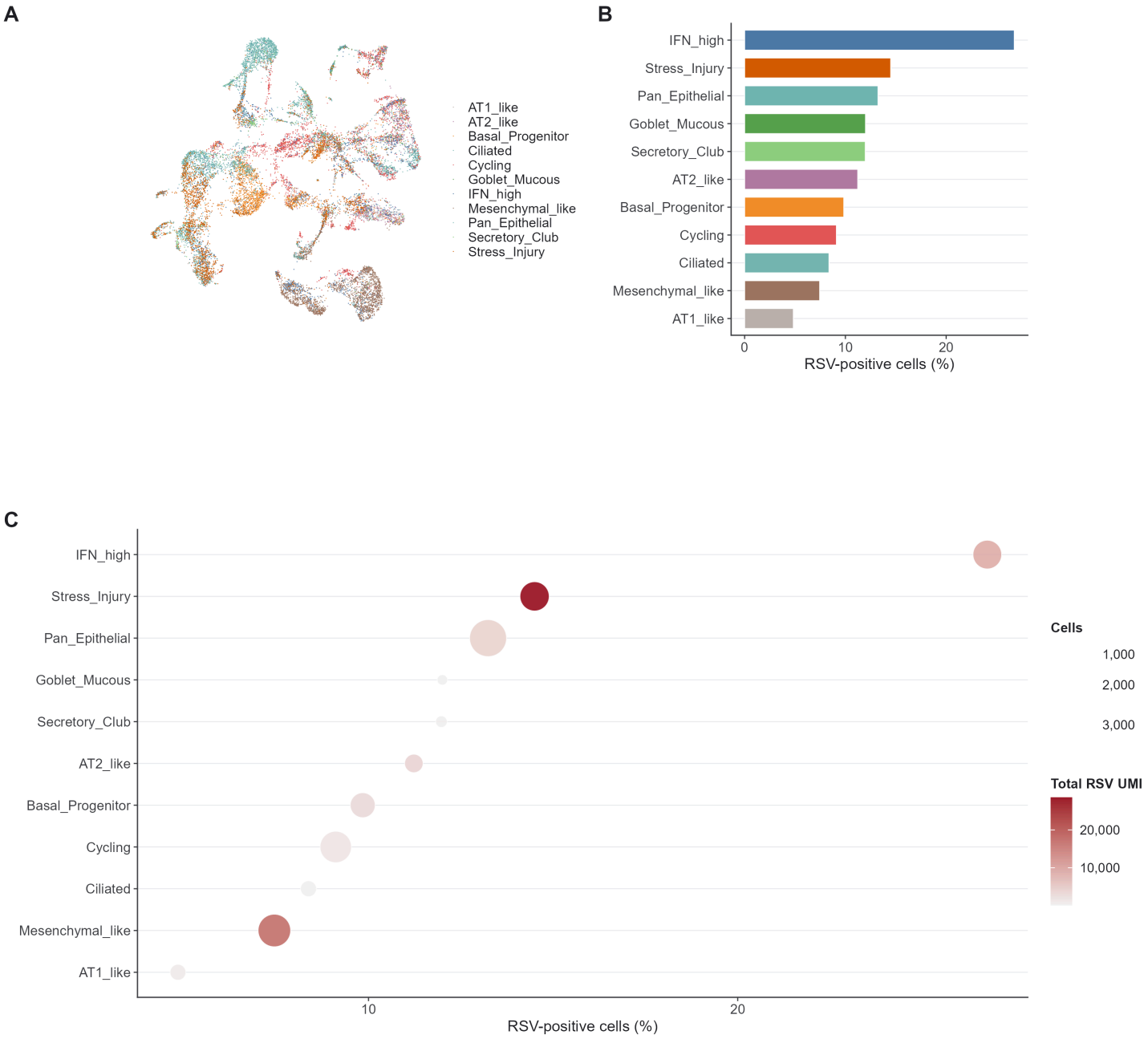


Figure 2

Refined cell states and RSV tropism. (A) Broad cell-state annotation of the RSV organoid atlas. (B) Refined epithelial and infection-associated state annotation. (C) Distribution of RSV RNA-positive cells across annotated states. (D) Enrichment of RSV signal in IFN-high and stress-injury states. The figure indicates that RSV transcript signal is not randomly distributed across transcriptional states.

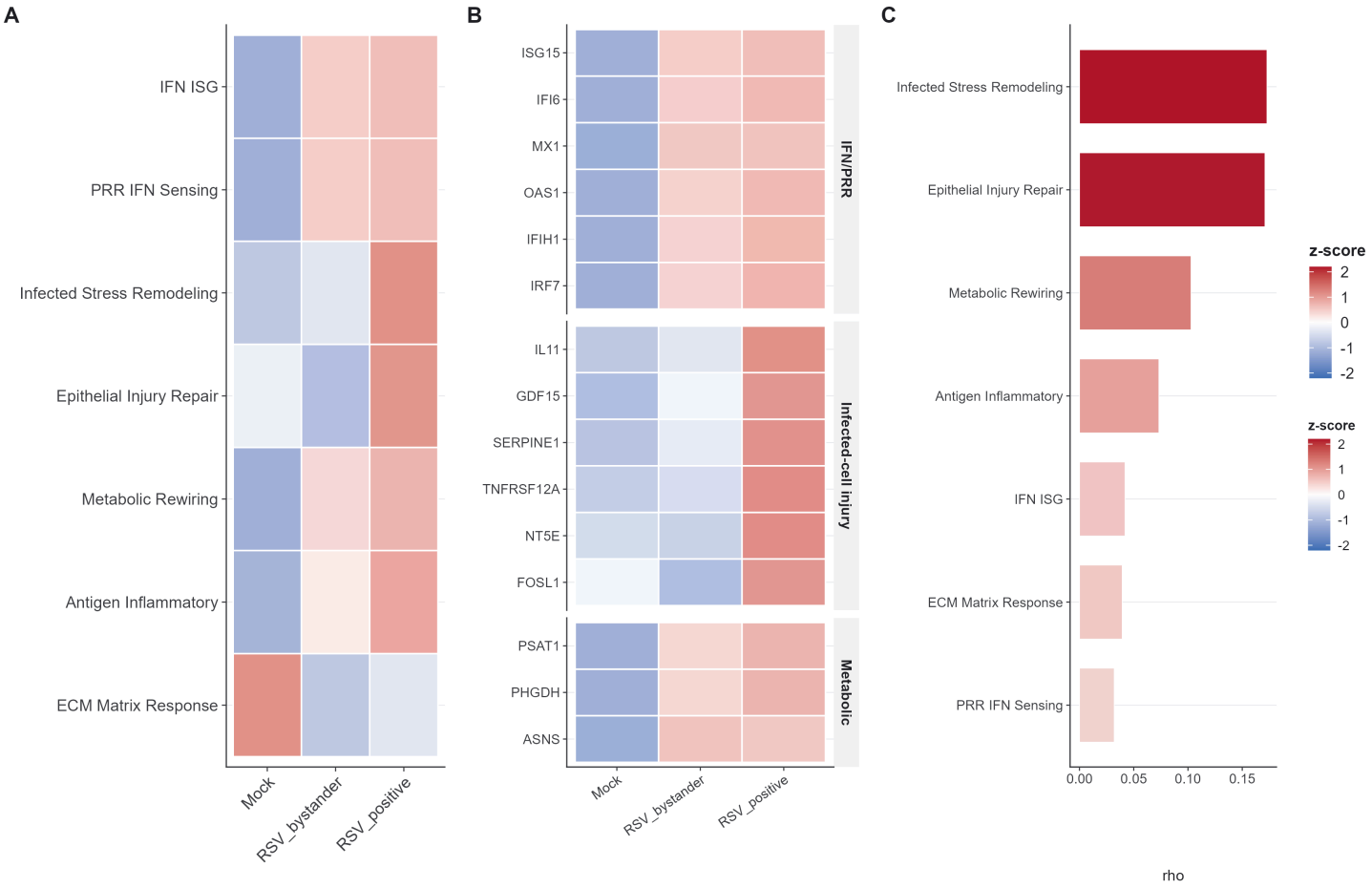


Figure 3

Mechanism programs in RSV-positive cells and bystanders. (A) Module-score comparison separates bystander antiviral sensing from RSV-positive injury-remodeling programs. (B) Representative interferon and pattern-recognition modules are enriched in bystander cells. (C) Stress-remodeling, epithelial injury-repair and metabolic-rewiring modules are higher in RSV transcript-positive cells. (D) Candidate-gene summaries link these modules to interpretable host-response nodes. Mean module differences and FDR-adjusted cell-level P values are reported in the source-data tables and Supplementary Data 2.

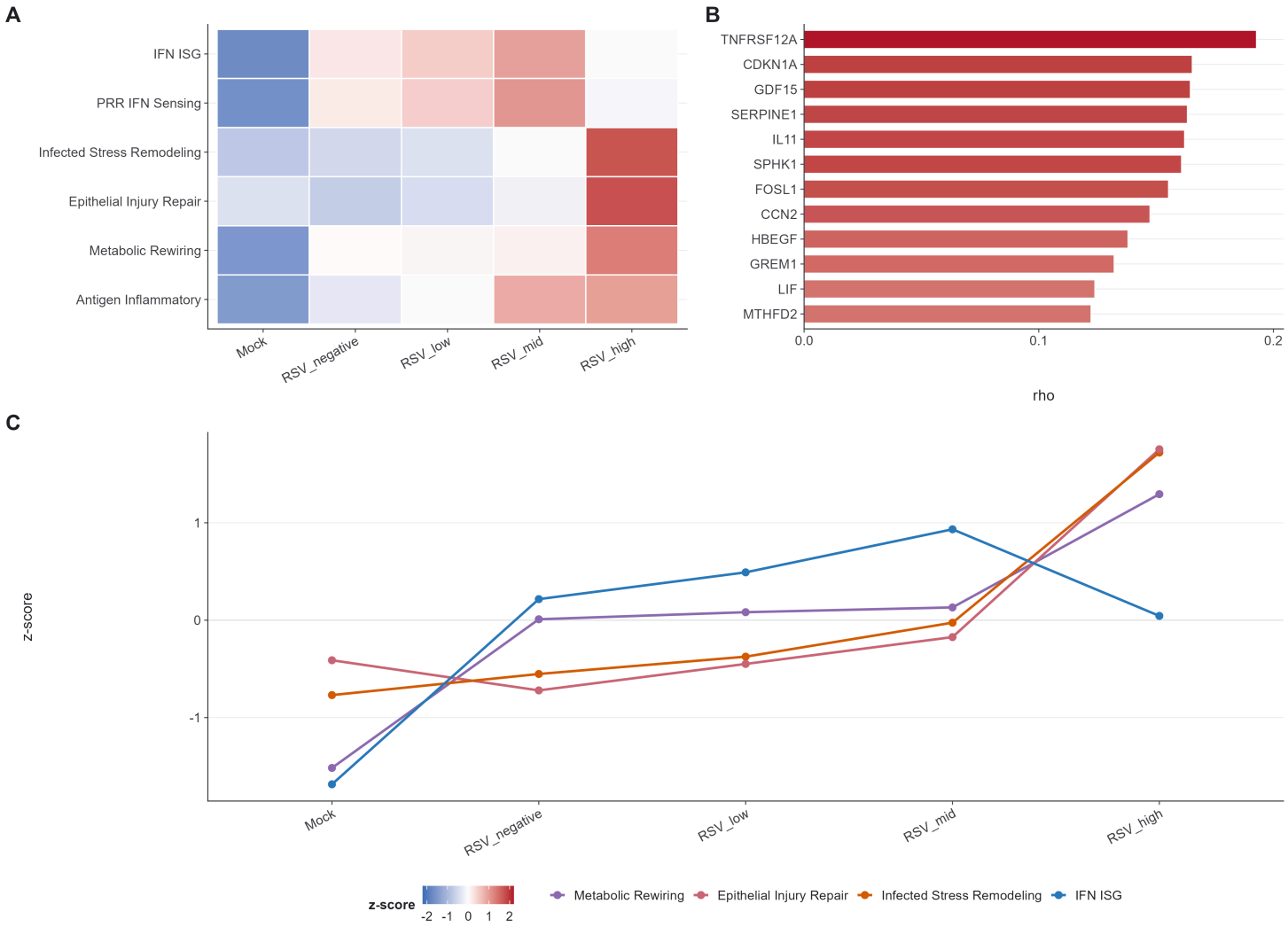


Figure 4

RSV viral-load gradient links viral burden to host injury programs. (A) RSV transcript-positive cells were stratified by viral UMI burden, with the RSV-high group defined by RSV UMI >=3. (B) Stress-remodeling scores increase across viral-load strata. (C) Epithelial injury-repair scores increase with viral burden. (D) Metabolic-rewiring scores show a weaker but concordant viral-load association. These panels support coupling between intracellular viral RNA and host injury-associated transcriptional states.

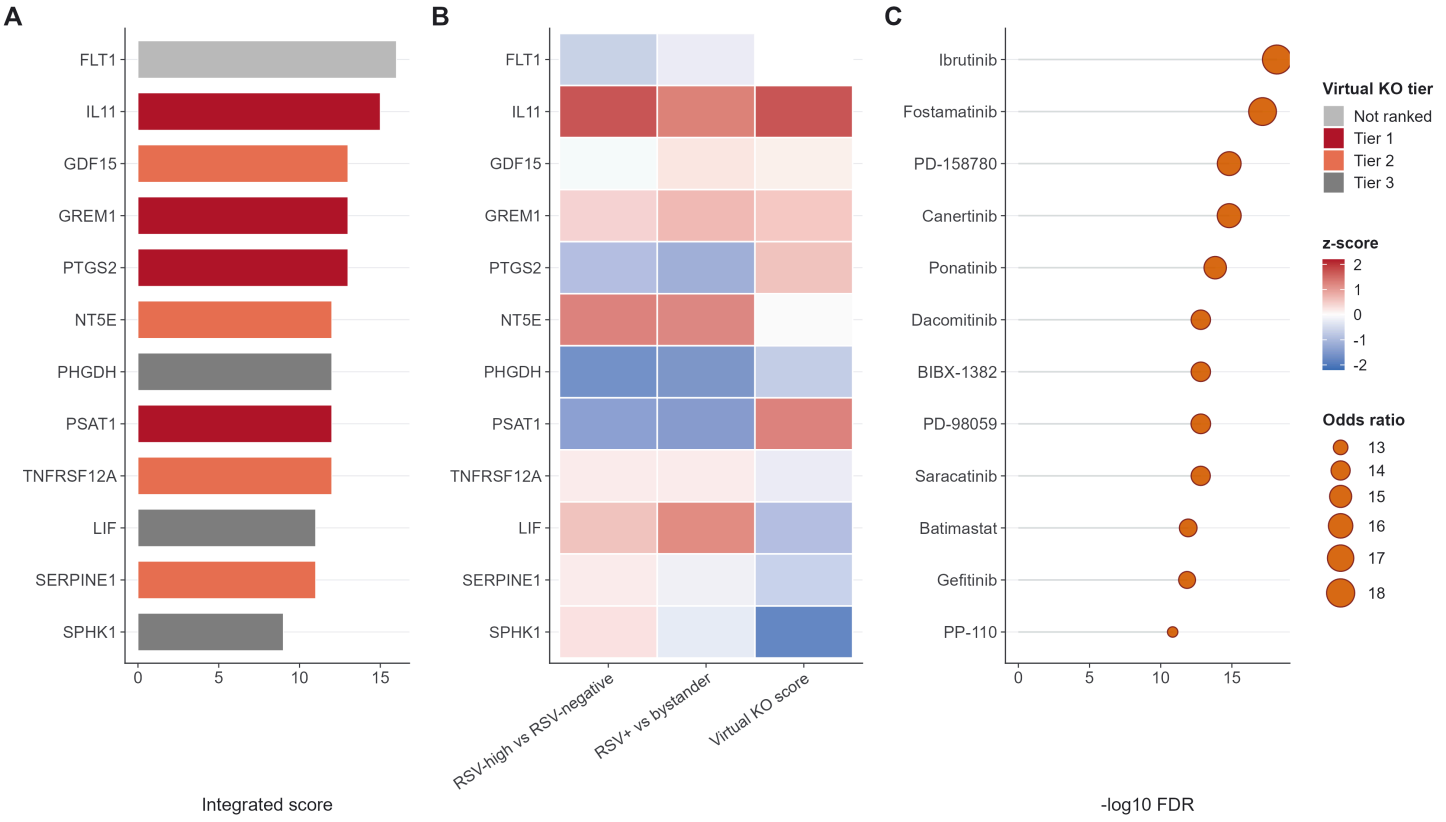


Figure 5

Integrated host targets and perturbation candidates. (A) Differential-expression and viral-load evidence prioritize candidate host-response genes. (B) Module membership links candidates to injury-repair, paracrine signaling, adenosine metabolism and metabolic adaptation. (C) Curated ligand-receptor scoring highlights remodeling axes such as IL11/IL6ST and TWEAK/TNFRSF12A. (D) Perturbation-reversal evidence adds an exploratory prioritization layer. The ligand-receptor panel uses an expression-based paracrine score and is best interpreted as mechanistic context, not formal cell-cell communication inference.

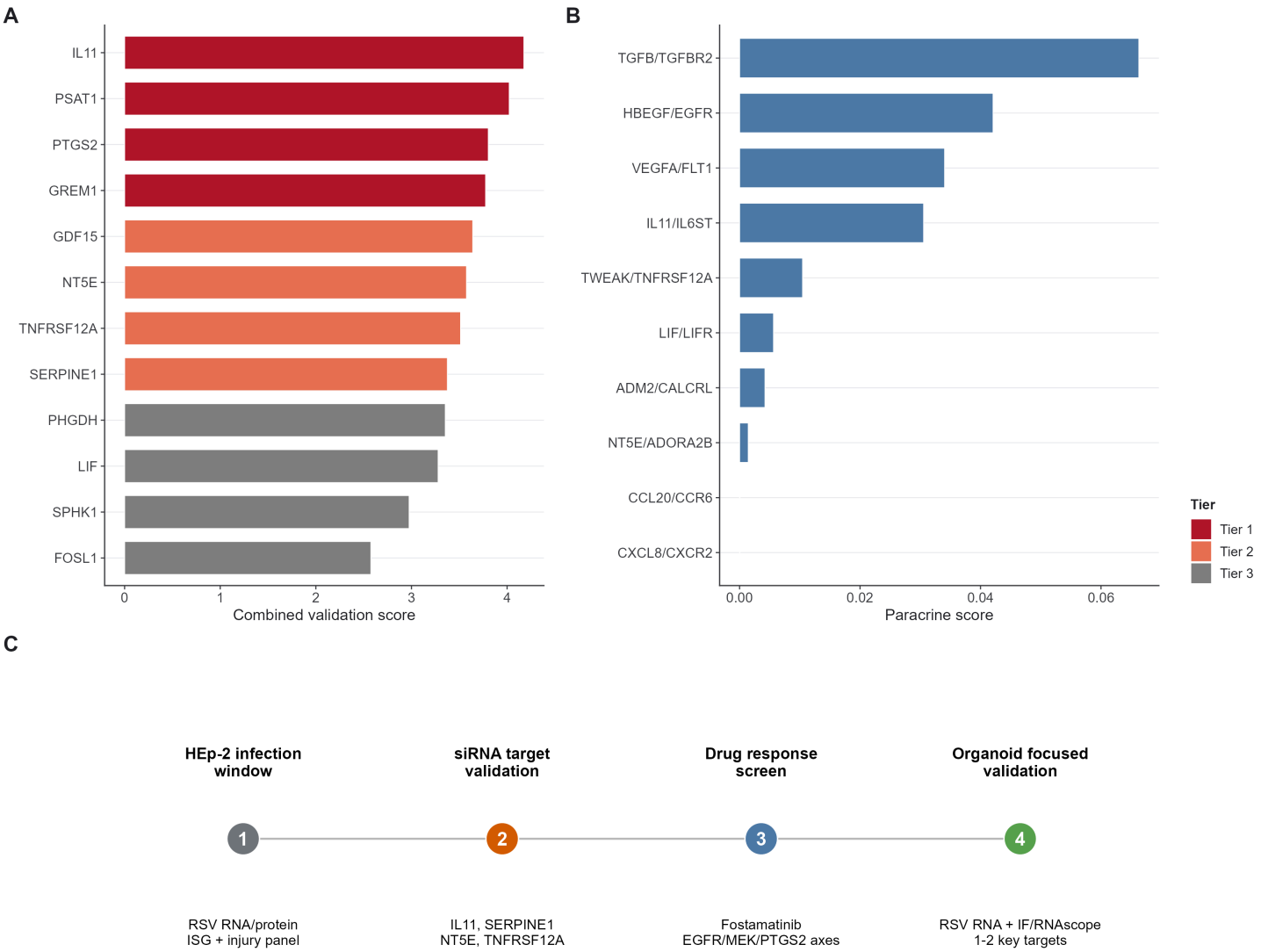


Figure 6

Exploratory virtual perturbation priority and follow-up roadmap. (A) Clean virtual knockout scoring ranks candidate genes by predicted network impact after removal of self-overlap effects. (B) Tier 1 and Tier 2 labels summarize computational priority groups. (C) Drug-signature reversal nominates perturbagens for follow-up screening. (D) The roadmap separates computational triage from future functional testing. Tier labels does not establish true knockout evidence or functional rescue.

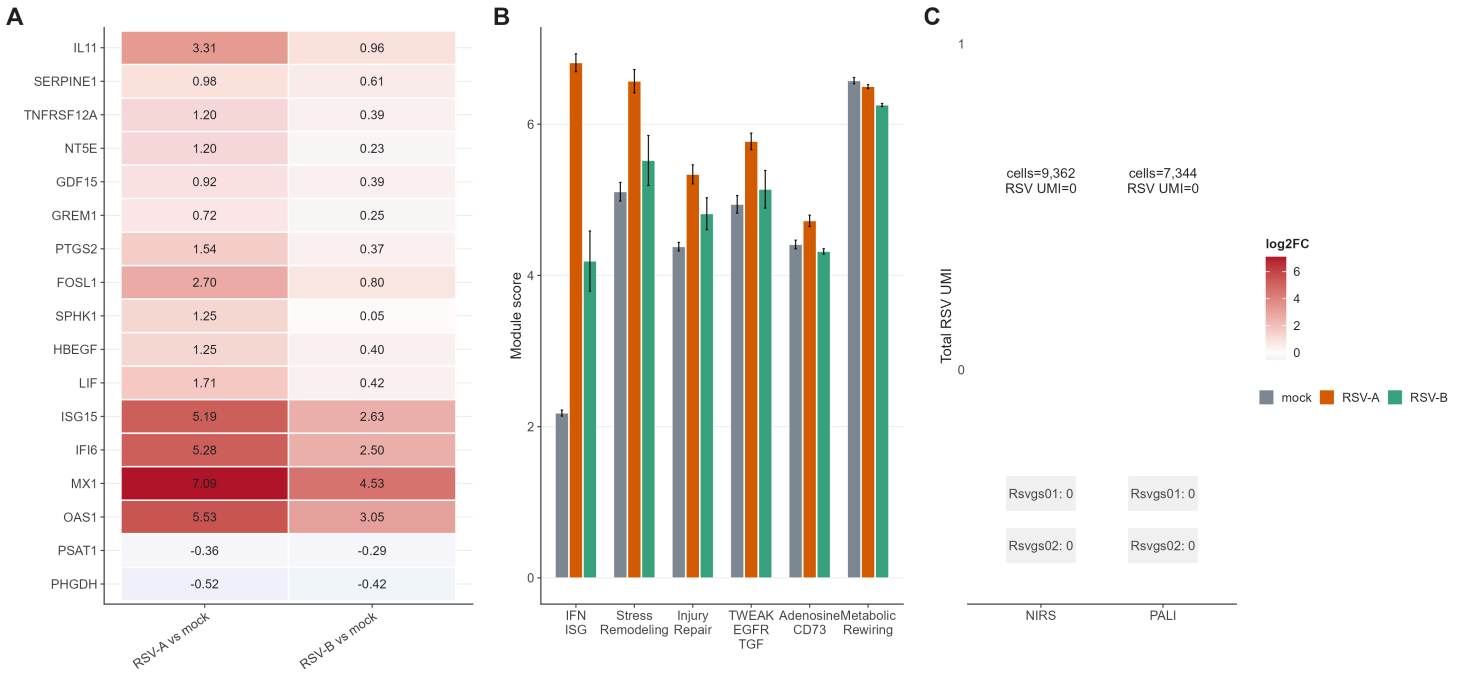


Figure 7

External organoid support for RSV injury-remodeling candidates. (A) GSE263272 bulk RNA-seq compares mock, RSV-A and RSV-B organoid conditions. (B) Candidate-gene expression in GSE263272 supports multiple discovery genes and modules. (C) GSE32139 primary airway epithelial data provide an additional epithelial reference for selected candidates. (D) GSE263271 antibody-treated scRNA-seq is shown only as a treatment-endpoint reference, not as an untreated infection support dataset.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation20260618.docx