Original Research

Title: AI-guided discovery of the invariant host response to viral pandemics

Authors: Debashis Sahoo[1-3†*], Gajanan D. Katkar[4*], Soni Khandelwal[1], Mahdi Behroozikhah[2], Amanraj Claire[4], Vanessa Castillo[4], Courtney Tindle[4], MacKenzie Fuller[4], Sahar Taheri[2], Thomas F. Rogers[5-6], Nathan Beutler[5], Sydney I. Ramirez[10, 11], Stephen A. Rawlings[11], Victor Pretorius[14], David M. Smith[11], Dennis R. Burton[5, 7-8], Laura E. Crotty Alexander[9], Jason Duran[15], Shane Crotty[10, 11], Jennifer M. Dan[10, 11], Soumita Das[11†] and Pradipta Ghosh[4,13†]

Affiliations:
[1]*Department of Pediatrics, University of California San Diego.*
[2]*Department of Computer Science and Engineering, Jacobs School of Engineering, University of California San Diego.*
[3]*Moores Cancer Center, University of California San Diego.*
[4]*Department of Cellular and Molecular Medicine, University of California San Diego.*
[5]*Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA.*
[6]*Division of Infectious Diseases, Department of Medicine, University of California, San Diego, La Jolla, CA 92037, USA.*
[7]*IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, CA 92037, USA.*
[8]*Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, CA 92037, USA.*
[9]*Pulmonary Critical Care Section, Veterans Affairs (VA) San Diego Healthcare System, La Jolla, California; Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of California San Diego (UCSD), La Jolla, California*
[10]*Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA, USA.*
[11]*Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA, USA.*
[12]*Department of Pathology, University of California San Diego.*
[13]*Medicine, University of California San Diego.*
[14]*Department of Surgery, University of California San Diego.*
[15]*Division of Cardiology, Department of Internal Medicine, UC San Diego Medical Center, La Jolla 92037*

*Equal contribution
† Co-Corresponding

Corresponding authors:
Debashis Sahoo, Ph.D.; Assistant Professor, Department of Pediatrics, University of California San Diego; 9500 Gilman Drive, MC 0730, Leichtag Building 132; La Jolla, CA 92093-0831. **Phone:** 858-246-1803: **Fax:** 858-246-0019: **Email:** dsahoo@ucsd.edu

Soumita Das, Ph.D.; Associate Professor, Department of Pathology, University of California San Diego; 9500 Gilman Drive, George E. Palade Bldg, Rm 256; La Jolla, CA 92093. **Phone**: 858-246-2062: **Email:** sodas@ucsd.edu

Pradipta Ghosh, M.D.; Professor, Departments of Medicine and Cell and Molecular Medicine, University of California San Diego; 9500 Gilman Drive (MC 0651), George E. Palade Bldg, Rm 232; La Jolla, CA 92093. **Phone**: 858-822-7633: **Email:** prghosh@ucsd.edu

ABSTRACT (197 words)

We sought to define the host immune response, a.k.a, the "cytokine storm" that has been implicated in fatal COVID-19 using an AI-based approach. Over 45,000 transcriptomic datasets of viral pandemics were analyzed to extract a 166-gene signature using ACE2 as a 'seed' gene; ACE2 was rationalized because it encodes the receptor that facilitates the entry of SARS-CoV-2 (the virus that causes COVID-19) into host cells. Surprisingly, this 166-gene signature was conserved in all *vi*ral *p*andemics, including COVID-19, and a subset of 20-genes classified disease severity, inspiring the nomenclatures *ViP* and *severe-ViP* signatures, respectively. The *ViP* signatures pinpointed a paradoxical phenomenon wherein lung epithelial and myeloid cells mount an IL15 cytokine storm, and epithelial and NK cell senescence and apoptosis determines severity/fatality. Precise therapeutic goals were formulated and subsequently validated in high-dose SARS-CoV-2-challenged hamsters using neutralizing antibodies that abrogate SARS-CoV-2•ACE2 engagement. IL15/IL15RA were elevated in the lungs of patients with fatal disease, and plasma levels of the cytokine tracked with disease severity. Thus, the *ViP* signatures provide a quantitative and qualitative framework for titrating the immune response in viral pandemics and may serve as a powerful unbiased tool to rapidly assess disease severity and vet candidate drugs.

One Sentence Summary: The host immune response in COVID-19.

Key Words:

- Artificial Intelligence/Machine Learning
- Boolean Equivalent Clusters
- Angiotensin Converting Enzyme (ACE)-2
- Coronavirus COVID-19
- Immune response
- Lung alveoli
- Natural Killer (NK) cells
- Interleukin 15 (IL15)

INTRODUCTION

As the rapidly unfolding COVID-19 pandemic claims its victims around the world, it has also inspired the scientific community to come up with solutions that have the potential to save lives. In the works are numerous investigational drugs at various phases of clinical trials, from rationalizing[1], to IRB approvals, recruitment and execution[2,3], all directed to meet an urgent and unmet need —i.e., ameliorate the severity of COVID-19 and reduce mortality.

Two obstacles make that task difficult—First, the pathophysiology of COVID-19 remains a mystery. The emerging reports generally agree that the disease has a very slow onset[4,5] and that those who succumb typically mount a 'cytokine storm'[4,6], i.e., an overzealous immune response. Despite being implicated as a cause of mortality and morbidity in COVID-19, we know virtually nothing about what constitutes (nature, extent) or contributes to (cell or origin) such an overzealous response. Consequently, treatment goals in COVID-19 have been formulated largely as a 'trial and error'-approach; this is reflected in the mixed results of the trials that have concluded[7]. Second, there is no established pre-clinical animal or human cell/organoid models for COVID-19; vetting the accuracy and/or the relevance of such models requires first an understanding of the host response in the disease.

We set out to define this aberrant host immune response in COVID-19 using machine learning tools that can look beyond interindividual variability to extract underlying gene expression patterns within multidimensional complex data. The approach was used across multiple cohorts of viral pandemics. The resultant pattern, i.e., signature, was subsequently exploited as a predictive model to navigate COVID-19. Findings not only pinpointed the precise nature of the cytokine storm, the culprit cell types and the organs, but also revealed disease pathophysiology, and helped formulate specific therapeutic goals for reducing disease severity. Key findings were validated in preclinical models of COVID-19 in Syrian hamsters and in the lungs and plasma of infected patients.

RESULTS AND DISCUSSION

An ACE2-centric Study Design

To identify and validate an invariant (universal) gene signature of host response in COVID-19, we mined more than 45,000 publicly available datasets of viral pandemics across three species (human, mouse and rats) (***Step 1***; **Fig 1**). Three relatively widely accepted facts shaped our approach using Angiotensin-converting enzyme 2 (ACE2) as 'seed' gene in our computational studies: *(i)* ACE2 is the most well-known portal for SARS-CoV-2 entry into the host cell[8,9]; its expression in cell lines correlates with the expression of innate immune genes [10] and susceptibility to SARS-CoV spike protein-driven entry[11,12], and its depletion in mice abrogates SARS-CoV infection[13]; *(ii)* ACE2 is a potent negative regulator of the renin–angiotensin aldosterone system (RAAS)[14]; without such restraint, the RAAS contributes to exuberant inflammation in the setting of infections[15]; and finally, *(iii)* although the mechanism through which ACE2 suppresses inflammatory response remains poorly understood, accumulating evidence indicates that infections perturb ACE2 activity, allowing for uncontrolled inflammation[16-24].

As ***Step 2*** (**Fig 1**), we validated the signature in several human and mouse datasets of viral pandemics, and a subset of genes was identified and validated as indicators of disease severity. The signatures were then validated in SARS-CoV-2-infected cells and tissues and to explore the nature, extent and cell of origin of host response in mild and fatal COVID-19.

As ***Step 3*** (**Fig 1**), the gene signatures were used to navigate the uncharted territory of COVID-19 and pinpoint immunopathologic mechanisms, which revealed the nature (IL15), source (airway epithelium), intensity (quantitative measure) and consequence (NK cell senescence) of the cytokine storm and helped objectively formulate precise therapeutic goals to reduce the severity of COVID-19.

As *Step 4* (**Fig 1**), the gene signature and the mechanism of action (IL15/IL15RA) was validated in lung tissues from SARS-CoV-2 challenged golden hamster using RNASeq and IHC. In addition, a precise therapeutic goals on SARS-CoV-2 was validated in the golden hamster model. The mechanism of action (IL15/IL15RA) was also validated by ELISA in plasma and IHC in lung tissues from UCSD COVID-19 cohort participants (Supplementary Table S4-5).

A Shared Host Response Signature in Respiratory Viral Pandemics

Because publicly available transcriptomic datasets from SARS-CoV-2-infected samples are still relatively few, any conclusion drawn from so few samples using any computational methodology is likely to lack robustness. We chose to use an informatics approach, i.e., Boolean Equivalent Correlated Clusters (BECC)[25], which can identify fundamental invariant (universal) gene expression relationships underlying any biological domain; in this case, we selected the biological domain of '*respiratory viral pandemics*'. BECC enables comparison of the normalized expression of two genes across all datasets by searching for two sparsely populated, diagonally opposite quadrants out of four possible quadrants (high-low and low-high), employing the BooleanNet algorithm[26]. There are six potential gene relationships assessed by BooleanNet: two symmetric (Equivalent and Opposite; **Fig 2A**) and four asymmetric[26]. Two genes are considered "Boolean Equivalent" if they are positively correlated with only high-high and low-low gene expression values. Two genes are considered "Boolean Opposite" if they are negatively correlated with only high-low and low-high gene expression values. Asymmetric Boolean implications result when there is only one sparsely populated quadrant. The BECC algorithm focuses exclusively on "Boolean Equivalent" relationships to identify potentially functionally related gene sets. Once identified, these invariant relationships have been shown to spur new fundamental discoveries[27,28], with translational potential[29], and most importantly, offer insights that aid the navigation of uncharted territories where nothing may be known[30,31].

We used GSE47963 [human airway epithelial (HAE) cultures with H1N1 and SARS-CoV infections; n = 438] as a 'test' dataset, which was comprised of human airway epithelial cell samples (HAE) infected *in vitro* with the causative agents of the 2009 'swine flu' (influenza A-H1N1; a triple recombination of human, avian, and swine influenza viruses[32-34]) and the 2002 Severe acute respiratory syndrome (SARS-CoV-1)[35] outbreaks (**Fig 2B**). These datasets were chosen now, and other datasets were prioritized later in the study, e.g., H5N1 (the causative agent of the avian flu in 2006-06[36]) and MERS-CoV (the causative agent of Middle East respiratory syndrome in 2012[37]) based upon the fact that they *all* contributed to outbreaks that are characterized by acute respiratory syndromes with high case-fatality rates[8].

ACE2 is used as a 'seed' to identify other genes that have 'Boolean Equivalent' and 'Boolean Opposite' relationships with ACE2. These genes were subsequently filtered using differential analysis on another dataset [GSE113211 (n = 118); **Fig 2B**] that profiled heterogeneous

immunophenotypes of children with viral bronchiolitis (confirmed positive for the virus in ~100% patients; of which 25 % were infected with Influenza/Para-Influenza and 14.8% with human CoV). Transcriptomes were analyzed in nasal mucosal scrapings (NMS) and PBMC samples taken during an acute visit (AV) and during a subsequent visit at convalescence (CV)[38]. 166 genes (**Table S1**; 1-1) retained the "Boolean Equivalent" relationship with ACE2 *and* their expression was downregulated during the convalescence visit. 26 genes (**Table S1;** 2-1) retained "Boolean Opposite" relationships with ACE2 *and* their expression were upregulated during the convalescence visit. All subsequent analyses were performed using the 166 –gene signature that had Boolean Equivalent relationship with ACE2 and that was down-regulated during a convalescent visit after acute viral bronchiolitis.

First the 166-gene signature was evaluated in the test dataset-- it was used to rank order the samples and test for phenotype classification using a receiver operating characteristic curve [ROC curve; the area under this curve (AUC) represents degree or measure of separability] and displayed such classification using violin plots (**Fig 2C-D**). The signature classified the uninfected *vs*. infected samples with reasonable accuracy in the setting of SARS-CoV-1 infection (ROC-AUC = 0.81, **Fig 2C**). It also classified perfectly in the setting of H1N1 infection (ROC-AUC = 1.00, **Fig 2D**). Good classification was observed between samples from the acute visit (AV) and convalescence visit (CV) in children (test dataset; GSE113211; **Fig 2E-F**, *left*), as well as two independent adult cohorts (validation datasets that were generated in two prospective studies[39,40]; **Fig 2G-H**). All the patients in these cohorts were infected with respiratory viruses; in one cohort, ~45% were documented infections with pandemic Influenza strains H1N1 and H3N2 (GSE68310), whereas 100% of the patients in the other were victims of the H1N1 pandemic of 2009 (GSE21802). Regardless of the heterogeneity of these validation cohorts, the classification score using the 166-gene signature remained strong in both datasets (ROC-AUC = 0.83 - 0.96). Findings indicate that the viral pandemic signature was conserved among numerous respiratory viral pandemics, and for that reason, we christened it the '*Vi*ral *P*andemic' (*ViP*) Signature.

The *ViP* Signature Defines the 'Cytokine Storm' in Viral Pandemics

Reactome analyses on the 166 genes showed that the signature was largely enriched for genes within the immune system pathways, e.g., interferon and cytokine signaling, cellular processes that are critical for an innate immune response such as the ER-phagosome pathway and antigen

processing and presentation, and finally the adaptive immune system (**Fig 3A-C**). In other words, the signature reflected a typical host immune response that is expected during any viral infection. This is not surprising because an overzealous host immune response, i.e., a 'cytokine storm' is shared among all respiratory viral pandemics (Influenza, avian and swine flu)[41] and severe COVID-19 patients who succumb to the disease[6]. However, there were 3 surprising factors: (i) This signature and reactome profile emerged using ACE2 as a 'seed' gene, which is not the receptor for influenza strains to enter into host cells. (ii) It is also noteworthy that despite filtration through two unrelated datasets (**Fig 2B**), one *in vitro* and another *in vivo*, and the reduction in the # of genes in the ACE2-equivalent cluster during such iterative refinement, the pathways/processes represented in the 166-gene cluster (**Table S1**; 1-2) remained virtually unchanged. (iii) The only cytokine/receptor pair that emerged in this 166-gene cluster was interleukin-15 (IL15/IL15RA; **Fig 3A, C**), indicating that transcripts of this cytokine are invariably equivalent with ACE2 expression across all datasets analyzed. Findings are in keeping with the well-established role of IL15 in both the pathogenesis[42] and the severity [43] of virus-induced lung injury. They are also consistent with the fact that IL15-/- mice are protected from lethal influenza[44].

Next we tested this 166-gene signature in numerous datasets of samples infected with viruses that have either caused pandemics in the past (SARS-CoV-1, MERS, Ebola, Zika, etc.) or continue to do so at present (Influenza A/B, HIV, HCV, etc). The signature perfectly classified uninfected and infected samples (ROC-AUC = 1.00; **Fig 3D**) in four humans (GSE56677, GSE45042, GSE17400, GSE30589) and two mouse SARS-CoV1 and MERS-CoV datasets (GSE19137, GSE52920). It also performed reasonably well in two other human and one mouse datasets (ROC-AUC ranging between 0.76-0.97; GSE37827, GSE33267, GSE50000; **Fig 3D**). Analysis of a time course of infection with SARS-CoV-1 (GSE33267; **Fig 3E**) revealed that classification of infected samples improved over time, beginning at 48 h and reaching perfection (ROC AUC = 1.00) at 60-72 h, which is consistent with epidemiologic findings in prior acute respiratory viral pandemics (SARS and MERS) have average incubation periods ranging ~2-7 days, which can sometimes last up to ~10-14 d. Among datasets with curated samples representing other viral pandemics that are neither respiratory nor acute, we found that classification scores for RNA viruses were significantly better compared to DNA viruses in *in vitro* systems (**Fig 3F** top, **Fig S1A**), especially for those that share clathrin-dependent endocytic methods to breach host cells (**Table S2**). However, the classification scores were indistinguishable between RNA and DNA

viruses in *in vivo* studies (**Fig 3F** bottom, **Fig S1A-B**). These results indicate that the 166-gene signature is shared among all viral pandemics, and not specific to respiratory viral pathogens.

Notably, the 166-gene host response signature was specific for viral infections; it performed poorly in classifying bacterial lung infections compared to viral infections ($p < 0.001$; **Fig 3G**). The signature also implicated the epithelial and myeloid cells, but not ECs and fibroblasts contribute to host immune response because the classification scores were better for airway epithelial cells (AE) and dendritic cells (DC) compared to fibroblasts (FI) and microvascular endothelial cells (ME) (ROC-AUC: 0.66, 0.82 vs 0.43, 0.37; **Fig 3H**). These scores correlated well with ACE2 expression in these different cell types ($p < 0.001$; **Fig 3H**), raising the possibility that viral entry through engagement of ACE2 and the induction of ACE2-equivalent host genes may be intertwined. That myeloid cells are major contributors to this signature was confirmed in five independent datasets; the 166-gene signature distinguished 'reactive' (M1-polarized) macrophages in them all (**Fig 3I**).

Together, these findings indicate that the ACE2-equivalent 166-gene signature is specific enough to distinguish between bacterial and viral pulmonary infections but is of broader relevance. The airway epithelial cells and cells of myeloid lineages (DCs and macrophages) appear to be major contributors to the *ViP* signature.

A 20-gene Subset within the *ViP* Signature Detects Disease Severity

To determine what constitutes 'severe/fatal' disease, we rank-ordered the 166 genes within the *ViP* signature for their ability to classify Influenza A/B-infected adult patients by clinical severity[45,46] (n = 154; **Fig 4A**). Severe disease was defined as intubation and mechanical ventilation due to poor oxygenation and/or death. A set of top 20 genes (**Fig 4A; Table S1, 3-1; Table S3**) was sufficient to classify healthy controls from infected patients (ROC-AUC = 1.00) as well as distinguish mild from severe disease with reasonable accuracy (ROC-AUC = 0.95) in the test cohort (**Fig 4B**). Reactome pathway analyses revealed that compared to the *ViP* signature, the 'severity'-related 20-gene cluster enriches a completely different set of cellular processes, i.e., DNA damage (especially induction of genes that are critical for base excision repair; BER), stress-induced senescence, neutrophil degranulation and changes in cell cycle (**Fig 4C, Fig S2**). We validated this signature side-by-side with the 166-gene *ViP* signature in three human datasets that included samples from mild vs. severe disease during the avian (H7N9), IAV (H3N1 and others)

and the swine (H1N1) flu viral pandemics (**Fig 4D,** *left*). Both the 166-gene *ViP* signature and the 20-gene severity signature performed similarly when it came to classifying control *vs*. mild disease, but the latter performed significantly better in classifying mild *vs*. severe disease, and did so consistently in both validation datasets (ROC-AUC ranging from 0.8-0.9; **Fig 4D,** *left*).

The severity signature performed well also in a large murine lung dataset in which mice were intranasally infected with non-lethal (NL, control), sub-lethal (SL, mild) and lethal (L, severe) doses of two different strains of H1N1 virus A; the Texas/36/91 (Tx91), which is non-lethal in C57Bl/6 mice and causes transient morbidity and compared against those infected with sublethal and lethal doses of the highly pathogenic Puerto Rico/8/34 (PR8), which causes ARDS and death in less than a week[47]. Harvested lungs were sorted into five different prospectively isolated cell subpopulations and analyzed by microarray (**Fig 4D,** *right*): alveolar macrophages, lymphocytes (BC, TC, NK), Ly6Chi mononuclear myeloid cells, neutrophils, CD45neg pulmonary epithelial cells. The 166-gene *ViP* signature distinguished the control *vs*. mild samples perfectly in all five cell types (ROC-AUC = 1.00; **Fig 4D,** *right*). The classification accuracy of the 20-gene severity signature, however, was most prominent in neutrophils (ROC-AUC = 1.00), followed by monocytes and macrophages (ROC-AUC = 0.9), and then epithelial cells (ROC-AUC = 0.8), but failed in lymphocytes. These findings suggest that the cells of the innate immune system are the primary contributors of disease severity.

We conclude that the 166-gene *ViP* signature that was initially built using *in vitro* infection datasets detects also the host immune response ('cytokine storm') in the complex *in vivo* systems where the response may be triggered by direct viral damage to the lung epithelium, but get propagated by feed-forward dysregulated immune response, both innate and adaptive. Surprisingly, this 166-gene *ViP* signature was not associated with disease severity; instead, severity-associated 20 genes that regulate stress and senescence-associated repression of protein expression and DNA damage (**Fig 4C**). DAVID GO analyses on the 20-gene signature indicated that 3 biological processes, e.g., transcriptional repression, apoptosis, and intermediates within the type I IFN (IFNγ signaling) pathway (**Fig 4E**) indicative of cellular distress, senescence/aging and death are the determinants of severity/fatality.

The *ViP* Signatures are Induced in the Lung Epithelial and Immune cells in COVID-19

We next tested the ability of the *ViP* signatures to distinguish between SARS-CoV-2-infected samples and uninfected controls in 3 independent datasets, 2 of which were datasets generated from cells infected *in vitro* (**Fig 5A-C**) and one that was generated from lung samples from a fatal case of COVID-19 (**Fig 5D**). The signature perfectly classified infected from uninfected samples in them all (ROC-AUC 1.00; **Fig 5A-B-D**); of the 166 genes, both IL15 and IL15RA were notably elevated in infected samples (**Fig 5A**). The 20-gene signature performed reasonably well in distinguishing infected from uninfected A549 cells (ROC-AUC = 0.87; **Fig 5E**), and the healthy from the COVID-19 lung sample (ROC-AUC = 1.00; **Fig 5G**), but not in airway cells (bronchial; ROC-AUC = 0.57; **Fig 5F**). In fact, the 166-gene and 20-gene signatures perfectly classified infected vs. uninfected samples in all *in vitro* cellular models of CoV-2 infection, regardless of the tissue/organ (**Fig 5H***; left, middle*). The signatures performed nearly perfectly (ROC-AUC = 0.90 - 1.00; **Fig 5H**, *right*) across all lung cell types from COVID-19 infected patients analyzed by single cell sequencing.

We next tested the ability of these signatures to distinguish mild *vs*. fatal COVID-19 in single cell sequencing datasets from patient-derived lung samples (**Fig 5I**). The 166-gene signature was able to distinguish control *vs*. mild infection most effectively in macrophages, airway epithelium, CD4+ T cells and NK cells (**Fig 5I**, *lower panel, lower row*) and mild *vs*. severe disease in the epithelium and in NK cells (**Fig 5I**, *lower panel, upper row*). The 20-gene signature not only performed well in classifying control *vs*. mild infection in the same 4 cell types as above but also in B cells and CD8+ T cells (**Fig 5I**, *upper panel, lower row*). However, the 20-gene severity signature continued to perform most optimally in the epithelium (ROC-AUC = 1.00) and in NK cells (**Fig 5I**, *upper panel, upper row*).

Together, these findings show that the 166-gene *ViP* signature seen in other respiratory viral pandemics is conserved also in COVID-19. The cytokine storm (166-genes, which included IL15/IL15RA; **Table S1**) was induced in multiple cell types; however, the 20-gene *ViP* signature of disease severity and fatality was most prominently induced in two cell types: (i) the airway epithelial cells, known producers of IL15 after viral infections[48],[49] and (ii) the NK cells which are known targets of physiologic as well as overzealous IL15 response[50,51].

Viral infection and IL15 Induce, and Flu Vaccine Attenuates the *ViP* Signatures in NK cells

We next asked how the *ViP* signatures are impacted when NK cells are exposed to virus-infected epithelial cells. NK cells are known to lyse influenza virus-infected cells by direct cytotoxicity and antibody-dependent cellular cytotoxicity (ADCC); enhancing such NK cell function has been shown to control influenza virus infections [52]. Clearance of other viruses (HIV-1, other retroviruses, etc.) and cancer immunotherapies also leverage such NK cell-dependent ADCC[53,54]. We analyzed a transcriptomic dataset (GSE115203)[55] generated from co-culture studies of human PBMCs (3 donors) with influenza (H1N1 Puerto Rico/08/1934)-infected airway epithelial cells (A549) (**Fig 6A**; *top*). PBMCs (from co-culture), or NK cells FACS-sorted from the PBMC were then analyzed by RNA Seq, and the study had confirmed NK cell ADCC responses were durably induced in this assay *via* type I IFN release from PBMCs. We found that both the 166- and 20-gene *ViP* signatures were induced in PBMCs and in NK cells sorted from the PBMCs (**Fig 6A**; *bottom left*), indicating that NK cells in these co-culture models were sufficient to capture the observed host immune response in patients with COVID-19.

To test the role of IL15 in the induction of *ViP* signatures, we leveraged three datasets—one that used recombinant IL15 (PBMCs; GSE77601), another that used anti-IL1Rβ mAb (mouse skin biopsies; GSE45551)[56], and a third study using the prototypic H3K27 demethylase inhibitor, GSK-J4; the latter was shown to inhibit NK cell effector cytokines in response to IL15 without impacting its cytotoxic killing activities (human, NK cells; GSE89484)[57]. Both *ViP* signatures were stimulated by IL15, but attenuated in the two other datasets where IL15's actions were blocked pharmacologically (**Fig 6A;** *right*). These findings indicate that IL15 could be necessary and sufficient to induce the *ViP* signatures.

Because two independent studies[58,59] recently showed that those vaccinated against influenza have lower odds of requiring intensive care, invasive ventilation and/or dying, we analyzed two transcriptomic datasets (GSE64655[60] and GSE133478[61]) in which PBMCs from subjects vaccinated with seasonal trivalent or quadrivalent influenza vaccine (TIV/QIV) were collected and analyzed for NK cell activation. The first study showed that the NK cells continued to demonstrate progressive attenuation of both the 166- and 20-gene signatures rapidly within 7 days (**Fig 6B**, *left*). The second study, in which the NK cell-enriched and depleted fractions collected pre- and post(30 d)-vaccination were tested for their response to re-stimulation with IL15 (low dose, 0.75 ng/ml, 18 h); such stimulation is known to enhance NK cell activity [62-64] and

promote viral clearance [65-67]. Both ViP signatures were attenuated post-vaccination in NK cell-enriched fractions, but not in depleted fractions (**Fig 6B**, *right*). Because such post-vaccination attenuation happened in the setting of experimentally confirmed[61] enhancement of overall NK cell response, we conclude that attenuation of *ViP* signatures among recipients of TIV could continue to offer protection during re-challenge. Because such protection is seen in NK-cell enriched, but not depleted fractions, we conclude that the protection is mediated primarily *via* the preservation of functional NK cells.

An IL15-storm Originating in the Lung Alveoli Determines the Severity of COVID-19

We next analyzed the *ViP* signatures in transcriptomic datasets generated from multiple organs at autopsy. Both the 166- and 20-gene *ViP* signatures were predominantly enhanced in one organ, the lungs (**Fig 6C-D**); and IL15/IL15RA were also elevated in the lungs (**Fig 6D**). These findings indicate that the 20- and 166-gene signatures go together, and suggest a plausible cause and effect relationship. For instance, severity-related cellular events (such as epithelial and NK cell senescence) occur in the milieu of the organ that mounts the highest IL15-predominant cytokine response, i.e., lungs. We also found that IL15 and its receptor IL15RA were significantly increased in severe COVID-19 lungs (**Fig 6E**). These findings predict that an overzealous IL15-predominant cytokine response is the most consistent finding in the most severe cases of COVID-19 and that the lung epithelium is the likely source of such a storm.

These predictions were validated in a cohort of symptomatic COVID-19 patients who presented to the UC San Diego Medical Center with varying disease severity, ranging from mild to fatal (see **Table S4**). Plasma ELISA studies revealed that IL15 levels were significantly elevated during the acute compared to the convalescent visit (**Fig 6F**), and in whom the clinical presentation was moderate-to-severe compared to those with mild disease (**Fig 6G**). A sub-group analysis confirmed that while gender or age did not have a significant impact on plasma IL15 levels independently, the aged male (> 40 y) cohort had a significantly higher IL15 level than the young males (**Fig 6H**; *left*). No such pattern was noted among females. These findings are consistent with the fact that the gender gap in COVID-19-related deaths widens markedly with age[68]. Lungs collected during autopsies from patients who succumbed to COVID-19 (see **Table S5**) further confirmed that lung epithelial cells, especially the alveolar type II pneomocytes and alveolar immune cell infiltrates express high levels of IL15 and its receptor, IL15RA (**Fig 7A-B**).

Taken together, these findings support the following model of the immunopathogenesis of COVID-19 (**Fig 7C**): Airway epithelial cells and cells of the myeloid lineage and other immune cells are the primary source of the 166-gene cytokine storm, of which, IL15 is a component. It is possible, that the primary target of IL15, i.e., NK cells, when exposed to this storm for a prolonged period undergo damage, stress-induced senescence and apoptosis. Our model is consistent with prior studies showing that the airway epithelial cells (especially bronchial) constitutively express the IL15 and IL15RA/B genes and that viral infections[49] and IFNγ induced the synthesis and secretion of IL15[49] and that prolonged and excessive stimulation with IL15 is known to induce NK cell exhaustion [50,51]. These findings are consistent with the emerging reports that NK cells are significantly exhausted and reduced in cases of severe COVID-19 infection[69,70] and that such reduction was seen as early as 3-6 days after the onset of symptoms[71]. We conclude that fatal COVID-19 is characterized by a paradoxical immune response, i.e., suppression of epithelial and NK cell functions (immunosuppression) in the setting of a cytokine storm (overzealous immune response).

The ViP Signatures Formulate Therapeutic Goals, Track Treatment Efficacy

Previously we showed that the attenuation of the *ViP* signature was 'associated' with the acquisition of natural convalescence in several respiratory viral pandemics (**Fig 2F-H**); we now asked if they could serve as a readout of 'causality' in the setting of therapeutic intervention. We analyzed interventional studies in the setting of other viral infections that shared the *ViP* signature, i.e., HCV, HIV, Zika and Ebola (**Fig 3H**; **Fig S1**; **Table S2**). The 166-gene *ViP* signature classified HCV-infected liver biopsies treated or not with directly acting anti-viral agents (DAAs) (**Fig 8A-C**) and HIV-infected samples treated or not with *a*nti-*r*etroviral *t*herapeutics (ART; ROC-AUC = 1.00; **Fig 8D**) with sufficient accuracy. In the case of Ebola, the *ViP* signature was somewhat effective in classifying crisis (i.e., acute) from convalescent PBMC samples (ROC AUC 0.64; **Fig 8E**), and previously described anti-Ebola therapeutic strategies (Topoisomerase depletion with siRNA[72] inhibited the signature in Ebola-infected alveolar epithelial cells (siTop; ROC AUC 1.00; **Fig 8F**)[72]. Finally, the *ViP* signature was accentuated in Zika infected human cortical neural progenitor cells (**Fig 8G**) and was effectively attenuated when these infected samples were treated with two investigational drugs that was found to be effective in inhibiting Zika infection. These findings suggest 'causality' between treatment (anti-viral therapies) and response (attenuation of

the *ViP* signature), and imply that attenuation of the 166-gene *ViP* signature is a desirable therapeutic goal.

We next sought to determine if the SARS-CoV-2 virus is responsible for the host response and validate the ability of the *ViP* signatures to track therapeutic response. To this end, we used SARS-CoV-2-neutralizing antibodies whose design was inspired by monoclonal antibodies (mAbs) isolated from from convalescent donors[73]. A specific isotype of this antibody, which binds to the receptor-binding domain (RBD-A) of SARS-CoV-2 spike protein in a fashion that precludes binding to host ACE2, was demonstrated as effective in preventing infection and weight loss symptoms, in cell based and *in vivo* hamster models of infection, respectively. We observed that SARS-CoV-2-challenged hamsters that were pre-treated with anti-CoV-2 antibody, but not the control Den3 antibody (see **Fig 8H** for experimental design) had 3 key findings: (i) they suppressed both the 166- and 20-*ViP* signatures that were otherwise induced in the infected lungs (**Fig 8I**); (ii) their lungs were protected from overwhelming immune cell infiltration and obliteration of alveolar space (**Fig 8J-K**); (iii) expression of IL15 and IL15 receptor was significantly reduced compared to what was observed in the infected lungs (**Fig 8J, L**). These results validate the ACE-centric computational approach for identifying the *ViP* signatures, i.e., when ACE2•virus engagement was disrupted using antibodies, the signatures were suppressed. The findings also indicate that the reversal of the signature and the IL15 storm could be used as a readout of therapeutic efficacy.

CONCLUSION

The major and unexpected finding in this work is that all viral pandemics (regardless of their acuity, causative virus, case fatality rates and clinical presentation) share a common fundamental host immune response. Summarized below are our three major findings.

First, we defined an invariant 166-gene host response --the so-called "cytokine storm"– that is surprisingly conserved among all viral pandemics. What was also unusual is that the signature emerged despite the rationalized use of ACE2 as a 'seed' gene to identify 'coronaviral infection-associated genes'. This suggests that while ACE2 may be the entry site for SARS-CoV-2, it is a prominently upregulated gene during host response to other viral infections. We conclude that the *ViP* signatures not just define the nature of the host immune response to viral pandemics, but also allows the tracking and quantification of the such response. *Second,* we define the precise nature of the cytokine storm and pinpoint the IL15 cytokine and its receptor, IL15RA as an

invariant component. We demonstrate that systemic levels of IL15 tracks with disease severity among patients and that the levels are notably elevated in the aged male (the predisposed age group in COVID-19, as per reports worldwide). Using a combination of single cell RNA Seq and human lung histology, we also pinpoint the lung epithelial and myeloid cells as the key contributors to the *ViP* signature, and more specifically, IL15/IL15RA. *Third*, we found that a subset of 20-gene 'severe' *ViP* signature, indicative of stress-induced senescence, transcriptional repression, DNA damage and apoptosis is also shared among various viral pandemics. In patients with COVID-19, this signature was seen in lung epithelial and NK cells, which is intriguing because airway epithelial cells is a prominent source and the NK cells are a major target of IL15. Thus, the *ViP* signatures begin to paint a picture of 'paradoxical immunosuppression' at the heart of fatal COVID-19, in which, the observed NK cell exhaustion/depletion in severe COVID-19[69-71] could be a consequence of an overzealous IL15 storm, leading to their senescense and apoptosis.

In closing, given that the emerging pandemic is still largely a mystery to us in terms of how it picks its victims, the *ViP* signature we define here provides a computational framework for navigation in otherwise uncharted territory. While it is expected that the signature will be more effective and accurate when it is iteratively filtered using emerging COVID-19 datasets, we provide evidence for its usefulness now in formulating therapeutic strategies and rapidly screening for therapeutics. Because the *ViP* signature of host response is seen also in other viral pandemics tested, findings may also be relevant also in navigating management strategies in those pandemics.

METHODS

Data Collection and Annotation: Publicly available microarray and RNA Seq databases were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) website[74-76]. Gene expression summarization was performed by normalizing Affymetrix platforms by RMA (Robust Multichip Average)[77,78] and RNASeq platforms by computing TPM (Transcripts Per Millions)[79,80] values whenever normalized data were not available in GEO.

Data Analysis: Boolean analysis and other statistical approaches are covered in detail in *Supplementary Online M*aterials. Briefly, the StepMiner algorithm[81], BooleanNet statistics[26], and BECC (Boolean Equivalent Correlated Clusters)[25] are used to perform Boolean analyses. Gene signature is computed by using a scaled linear combination of gene expression values which is used to classify sample categories and the performance of the multi-class classification is measured by ROC AUC (Receiver Operating Characteristics Area Under The Curve) values. A color-coded bar plot is combined with a violin plot to visualize the gene signature-based classification and distribution of the gene signature score. Bubble plots of ROC-AUC values (radius of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification based on the 20 gene severe ViP signature and 166 gene ViP signature is visualized side by side.

Validation studies

Plasma ELISA studies on COVID-19 patients: Blood from COVID-19 donors was either obtained at a UC San Diego Health clinic or recruited at the La Jolla Institute under active IRB protocols. Levels of IL15 cytokine was estimated using ELISA MAX Deluxe set (Biolegend Cat. No. 435104) according to the manufacturer's recommended protocol.

Autopsy and biopsies of lungs: The detailed process of collecting and processing lung specimens from the COVID 19 positive human subjects is available in the supplementary methods.

Hamster study with anti-CoV-2 therapy: Lung samples from 8-week old Syrian hamsters were generated exactly previously published study[11]. See Supplementary methods for details regarding IHC protocols.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes for Health (NIH) grants R00-CA151673 and R01-GM138385 (to DS) and R01-AI141630, CA100768 and CA160911 (to P.G), R01DK107585 (SD) and R01-AI155696 (to P.G, D.S and S.D), UCOP-RGPO (R00RG2628 & R00RG2642 to P.G, D.S and S.D), The Sanford Stem Cell Clinical Center at UCSD (to P.G, D.S and S.D) and U19-AI142742 (to S.C, CCHI: Cooperative Centers for Human Immunology), and LJI Institutional Funds (to S.C). Dr. Crotty Alexander's salary was supported in part by the VA San Diego Healthcare System. The authors would like to thank Rachel White (RW) and Jen Bigbee (JB), who assisted with both thoracotomies/biopsies and helped set up and design the autopsy study.

AUTHOR CONTRIBUTIONS

Role	Authors
Conceptualization; Funding acquisition; Supervision; Project administration	D.S., S.D., P.G.
Investigation; Methodology; Data curation, Formal Analysis • Computational modeling (D.S) • Data curation and analysis (D.S., P.G., S.D. S.T., M.B., G.D.K) • Hamster studies (C.T, M.R.F., S.K., A.C., V.C.) • Rapid Autopsy of Lungs (S.K, A.C., V.C.) • Plasma ELISA (G.D.K.) • Essential resources (animal and patient samples, tissues, etc) (T.F.R, N.B, D.R.B, J.D., L.C.A, S.C., J.D., S.I.R, V.P., S.A.R, D.M.S)	D.S., S.T, G.D.K., T.F.R, N.B, D.R.B, C.T, M.R.F., J.D., L.C.A, S.I.R, V.P, S.A.R, D.M.S, S.K, M.B., S.D., P.G.
Funding acquisition	D.S., S.D., P.G.
Software	D.S.
Visualization; Writing – original draft; Writing – review & editing	D.S., P.G., S.D.

Competing interests: The authors declare no competing interests.

Data and materials availability:

All data is available in the main text or the supplementary materials.

Supplementary Materials: Includes

- Detailed Materials and Methods
- Supplementary Text- **n/a**
- Figures **S1, S2**
- Tables **S1-5**
- External Databases **- None**
- References **(*1-15)***

REFERENCES

1. Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. Lancet 2020;395:e30-e1.

2. Amon A. The spindle checkpoint. Curr Opin Genet Dev 1999;9:69-75.

3. Bertoli C, Skotheim JM, de Bruin RA. Control of cell cycle transcription during G1 and S phases. Nat Rev Mol Cell Biol 2013;14:518-28.

4. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 2020;382:1199-207.

5. Fauci AS, Lane HC, Redfield RR. Covid-19 - Navigating the Uncharted. N Engl J Med 2020;382:1268-9.

6. Mehta P, McAuley DF, Brown M, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet 2020;395:1033-4.

7. Mixed results from recently completed Covid-19 clinical trials. https://www.clinicaltrialsarena.com, 2020. (Accessed 4/4/2020, 2020, at https://www.clinicaltrialsarena.com/comment/covid-19-clinical-trials-results/.)

8. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 2020.

9. Tai W, He L, Zhang X, et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. Cell Mol Immunol 2020.

10. Sungnak W, Huang N, Becavin C, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat Med 2020;26:681-7.

11. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci U S A 2005;102:7988-93.

12. Nie Y, Wang P, Shi X, et al. Highly infectious SARS-CoV pseudotyped virus reveals the cell tropism and its correlation with receptor expression. Biochem Biophys Res Commun 2004;321:994-1000.

13. Kuba K, Imai Y, Rao S, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. Nat Med 2005;11:875-9.

14. Rodrigues Prestes TR, Rocha NP, Miranda AS, Teixeira AL, Simoes ESAC. The Anti-Inflammatory Potential of ACE2/Angiotensin-(1-7)/Mas Receptor Axis: Evidence from Basic and Clinical Research. Curr Drug Targets 2017;18:1301-13.

15. Hanafy S, Tavasoli M, Jamali F. Inflammation alters angiotensin converting enzymes (ACE and ACE-2) balance in rat heart. Inflammation 2011;34:609-13.

16. Sodhi CP, Nguyen J, Yamaguchi Y, et al. A Dynamic Variation of Pulmonary ACE2 Is Required to Modulate Neutrophilic Inflammation in Response to Pseudomonas aeruginosa Lung Infection in Mice. J Immunol 2019;203:3000-12.

17. Jia H. Pulmonary Angiotensin-Converting Enzyme 2 (ACE2) and Inflammatory Lung Disease. Shock 2016;46:239-48.

18. Kuba K, Imai Y, Penninger JM. Multiple functions of angiotensin-converting enzyme 2 and its relevance in cardiovascular diseases. Circ J 2013;77:301-8.

19. Jiang F, Yang J, Zhang Y, et al. Angiotensin-converting enzyme 2 and angiotensin 1-7: novel therapeutic targets. Nat Rev Cardiol 2014;11:413-26.

20. Sodhi CP, Wohlford-Lenane C, Yamaguchi Y, et al. Attenuation of pulmonary ACE2 activity impairs inactivation of des-Arg(9) bradykinin/BKB1R axis and facilitates LPS-induced neutrophil infiltration. Am J Physiol Lung Cell Mol Physiol 2018;314:L17-L31.

21. Simoes e Silva AC, Silveira KD, Ferreira AJ, Teixeira MM. ACE2, angiotensin-(1-7) and Mas receptor axis in inflammation and fibrosis. Br J Pharmacol 2013;169:477-92.

22. Liu X, Yang N, Tang J, et al. Downregulation of angiotensin-converting enzyme 2 by the neuraminidase protein of influenza A (H1N1) virus. Virus Res 2014;185:64-71.

23. Imai Y, Kuba K, Rao S, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature 2005;436:112-6.

24. Gaddam RR, Chambers S, Bhatia M. ACE and ACE2 in inflammation: a tale of two enzymes. Inflamm Allergy Drug Targets 2014;13:224-34.

25. Dabydeen SA, Desai A, Sahoo D. Unbiased Boolean analysis of public gene expression data for cell cycle gene identification. Mol Biol Cell 2019;30:1770-9.

26. Sahoo D, Dill DL, Gentles AJ, Tibshirani R, Plevritis SK. Boolean implication networks derived from large scale, whole genome microarray datasets. Genome Biol 2008;9:R157.

27. Inlay MA, Bhattacharya D, Sahoo D, et al. Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. Genes Dev 2009;23:2376-81.

28. Dalerba P, Kalisky T, Sahoo D, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat Biotechnol 2011;29:1120-7.

29. Dalerba P, Sahoo D, Paik S, et al. CDX2 as a Prognostic Biomarker in Stage II and Stage III Colon Cancer. N Engl J Med 2016;374:211-22.

30. Volkmer JP, Sahoo D, Chin RK, et al. Three differentiation states risk-stratify bladder cancer into distinct subtypes. Proc Natl Acad Sci U S A 2012;109:2078-83.

31. Sahoo D, Wei W, Auman H, et al. Boolean analysis identifies CD38 as a biomarker of aggressive localized prostate cancer. Oncotarget 2018;9:6550-61.

32. Vasa CV. State of the Globe: The H1N1 Threat Continues to Loom the Planet. J Glob Infect Dis 2013;5:91-2.

33. van der Meer FJ, Orsel K, Barkema HW. The new influenza A H1N1 virus: balancing on the interface of humans and animals. Can Vet J 2010;51:56-62.

34. Baden LR, Drazen JM, Kritek PA, Curfman GD, Morrissey S, Campion EW. H1N1 influenza A disease--information for health professionals. N Engl J Med 2009;360:2666-7.

35. Zhong NS, Zheng BJ, Li YM, et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet 2003;362:1353-8.

36. Beigel JH, Farrar J, Han AM, et al. Avian influenza A (H5N1) infection in humans. N Engl J Med 2005;353:1374-85.

37. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814-20.

38. Jones AC, Anderson D, Galbraith S, et al. Personalized Transcriptomics Reveals Heterogeneous Immunophenotypes in Children with Viral Bronchiolitis. Am J Respir Crit Care Med 2019;199:1537-49.

39. Bermejo-Martin JF, Martin-Loeches I, Rello J, et al. Host adaptive immunity deficiency in severe pandemic influenza. Crit Care 2010;14:R167.

40. Zhai Y, Franco LM, Atmar RL, et al. Host Transcriptional Response to Influenza and Other Acute Respiratory Viral Infections--A Prospective Cohort Study. PLoS Pathog 2015;11:e1004869.

41. Peiris JS, Hui KP, Yen HL. Host response to influenza virus: protection versus immunopathology. Curr Opin Immunol 2010;22:475-81.

42. Nakamura R, Maeda N, Shibata K, Yamada H, Kase T, Yoshikai Y. Interleukin-15 is critical in the pathogenesis of influenza a virus-induced acute lung injury. J Virol 2010;84:5574-82.

43. Leahy TR, McManus R, Doherty DG, et al. Interleukin-15 is associated with disease severity in viral bronchiolitis. Eur Respir J 2016;47:212-22.

44. Abdul-Careem MF, Mian MF, Yue G, et al. Critical role of natural killer cells in lung immunopathology during influenza infection in mice. J Infect Dis 2012;206:167-77.

45. Tang BM, Shojaei M, Teoh S, et al. Neutrophils-related host factors associated with severe disease and fatality in patients with influenza infection. Nat Commun 2019;10:3422.

46. Zerbib Y, Jenkins EK, Shojaei M, et al. Pathway mapping of leukocyte transcriptome in influenza patients reveals distinct pathogenic mechanisms associated with progression to severe infection. BMC Med Genomics 2020;13:28.

47. Brandes M, Klauschen F, Kuchen S, Germain RN. A systems analysis identifies a feedforward inflammatory circuit leading to lethal influenza infection. Cell 2013;154:197-212.

48. Zdrenghea MT, Telcian AG, Laza-Stanca V, et al. RSV infection modulates IL-15 production and MICA levels in respiratory epithelial cells. Eur Respir J 2012;39:712-20.

49. Ge N, Nishioka Y, Nakamura Y, et al. Synthesis and secretion of interleukin-15 by freshly isolated human bronchial epithelial cells. Int Arch Allergy Immunol 2004;135:235-42.

50. Felices M, Lenvik AJ, McElmurry R, et al. Continuous treatment with IL-15 exhausts human NK cells via a metabolic defect. JCI Insight 2018;3.

51. Judge SJ, Murphy WJ, Canter RJ. Characterizing the Dysfunctional NK Cell: Assessing the Clinical Relevance of Exhaustion, Anergy, and Senescence. Front Cell Infect Microbiol 2020;10:49.

52. NCI-Navy Medical Oncology Branch cell line supplement. J Cell Biochem Suppl 1996;24:1-291.

53. Parsons MS, Richard J, Lee WS, et al. NKG2D Acts as a Co-Receptor for Natural Killer Cell-Mediated Anti-HIV-1 Antibody-Dependent Cellular Cytotoxicity. AIDS Res Hum Retroviruses 2016;32:1089-96.

54. Pegram HJ, Andrews DM, Smyth MJ, Darcy PK, Kershaw MH. Activating and inhibitory receptors of natural killer cells. Immunol Cell Biol 2011;89:216-24.

55. Jegaskanda S, Vanderven HA, Tan HX, et al. Influenza Virus Infection Enhances Antibody-Mediated NK Cell Functions via Type I Interferon-Dependent Pathways. J Virol 2019;93.

56. Xing L, Dai Z, Jabbari A, et al. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. Nat Med 2014;20:1043-9.

57. Cribbs A, Hookway ES, Wells G, et al. Inhibition of histone H3K27 demethylases selectively modulates inflammatory phenotypes of natural killer cells. J Biol Chem 2018;293:2422-37.

58. Fink G, Orlova-Fink N, Schindler T, et al. Inactivated trivalent influenza vaccine is associated with lower mortality among Covid-19 patients in Brazil. medRxiv 2020:2020.06.29.20142505.

59. Zanettini C, Omar M, Dinalankara W, et al. Influenza Vaccination and COVID19 Mortality in the USA. medRxiv 2020:2020.06.24.20129817.

60. Hoek KL, Samir P, Howard LM, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PLoS One 2015;10:e0118528.

61. Wagstaffe HR, Pickering H, Houghton J, et al. Influenza Vaccination Primes Human Myeloid Cell Cytokine Secretion and NK Cell Function. J Immunol 2019;203:1609-18.

62. Carson WE, Fehniger TA, Haldar S, et al. A potential role for interleukin-15 in the regulation of human natural killer cell survival. J Clin Invest 1997;99:937-43.

63. Carson WE, Giri JG, Lindemann MJ, et al. Interleukin (IL) 15 is a novel cytokine that activates human natural killer cells via components of the IL-2 receptor. J Exp Med 1994;180:1395-403.

64. Carson WE, Ross ME, Baiocchi RA, et al. Endogenous production of interleukin 15 by activated human monocytes is critical for optimal production of interferon-gamma by natural killer cells in vitro. J Clin Invest 1995;96:2578-82.

65. Garrido C, Abad-Fernandez M, Tuyishime M, et al. Interleukin-15-Stimulated Natural Killer Cells Clear HIV-1-Infected Cells following Latency Reversal Ex Vivo. J Virol 2018;92.

66. van Erp EA, van Kampen MR, van Kasteren PB, de Wit J. Viral Infection of Human Natural Killer Cells. Viruses 2019;11.

67. Verbist KC, Klonowski KD. Functions of IL-15 in anti-viral immunity: multiplicity and variety. Cytokine 2012;59:467-78.

68. Bhopal SS, Bhopal R. Sex differential in COVID-19 mortality varies markedly by age. Lancet 2020;396:532-3.

69. Yaqinuddin A, Kashir J. Innate immunity in COVID-19 patients mediated by NKG2A receptors, and potential treatment using Monalizumab, Cholroquine, and antiviral agents. Med Hypotheses 2020;140:109777.

70. Zheng M, Gao Y, Wang G, et al. Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell Mol Immunol 2020;17:533-5.

71. Leng Z, Zhu R, Hou W, et al. Transplantation of ACE2(-) Mesenchymal Stem Cells Improves the Outcome of Patients with COVID-19 Pneumonia. Aging Dis 2020;11:216-28.

72. Rialdi A, Campisi L, Zhao N, et al. Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation. Science 2016;352:aad7993.

73. Rogers TF, Zhao F, Huang D, et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science 2020;369:956-63.

74. Barrett T, Suzek TO, Troup DB, et al. NCBI GEO: mining millions of expression profiles--database and tools. Nucleic Acids Res 2005;33:D562-6.

75. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013;41:D991-5.

76. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207-10.

77. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.

78. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-64.

79. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:323.

80. Pachter L. Models for transcript quantification from RNA-Seq. arXiv e-prints2011.

81. Sahoo D, Dill DL, Tibshirani R, Plevritis SK. Extracting binary signals from microarray time-course data. Nucleic Acids Res 2007;35:3705-12.

FIGURE LEGENDS

Figure 1. Study design.

(*From top to bottom*) *Step 1*: A database containing > 45,000 human, mouse and rat gene–expression data was mined to identify and validate an invariant signature for host response to viral pandemic (ViP) infection. ACE2, the portal for SARS-CoV-2 entry/uptake, was used as a 'seed' gene and Boolean Equivalent Correlated Clusters (BECC) was used as the computational method to identify gene clusters that share invariant relationships with ACE2. Once defined, these gene clusters (a.k.a., '*ViP signature*') were subsequently validated across numerous human and murine models of pandemic viral infection. *Step 2:* A subset of 20 genes from the ViP signature was selected that was strongly associated with severity of viral infection. These genes were validated in other cohorts to establish the 'Severe' ViP signature. Both 166- and 20-gene ViP signatures were validated on COVID-19 datasets. *Step 3*: Cross-validation studies in numerous other datasets helped- (i) define the nature (ii) and source of the cytokine storm in COVID-19, (iii) gain insights into the immunopathology of fatal disease, and (iv) set precise therapeutic goals. *Step 4*: Findings in step 3 were validated in hamsters and in a cohort of COVID-19 patients.

Figure 2. Identification and validation of an invariant ACE2-centric signature of host response to viral infections.

(A) Computational approach to identify Boolean Equivalent and Opposite relationships. Number of samples in all four quadrants are used to compute two parameters (S, p). $S > 5$ and $p < 0.05$ is used to identify sparse quadrant. Equivalent relationships are discovered when top-left and bottom-right quadrants are sparse (*left*). Opposite relationships are discovered when top-right and bottom-left quadrants are sparse (*right*). **(B)** Schematic displaying the key computational steps and findings leading to the identification of the 166-gene host response signature using ACE2 as a 'seed' gene. See also **Table S1**. **(C)** Bar and violin plots displaying sample rank order (i.e., classification) of SARS-CoV-1-infected samples and distribution of the 166 gene-based signature in the test dataset (GSE47963, *in vitro* infections of human airway epithelial cells). ROC-AUC values of infected samples classifications are shown below each bar plot unless otherwise stated. **(D)** Analysis of H1N1-infected samples compared to uninfected controls using the 166-gene signature like C. **(E)** Classification of patient samples used in datasets **F-H** based on their time of

collection either during 'Acute Visit' (AV) in the setting of an acute respiratory viral infection and 'Convalescence Visit' (CV) after recovery. **(F)** Analysis of PBMC samples from children (GSE113211, left) and nasal mucosal scrapings (NMS, GSE113211, right). **(G)** Analysis of peripheral blood from adults (GSE68310). **(H)** Analysis of patient samples collected during the swine flu pandemic (GSE21802).

Figure 3. Validation of the ViP signatures in global pandemic viral infections.

(A) Heatmap of the 166-gene signature on test dataset (GSE47963, *in vitro* infections of human airway epithelial cells). Genes that are involved in cytokine signaling in immune system are highlighted in the left. **(B)** ReacFoam analysis on the 166-gene signature that visualizes genome-wide pathway analysis based on Voronoi tessellation. **(C)** Reactome pathway analysis of the 166 genes in the Vip signature. **(D)** ViP signature-based classification of CoV-infected samples (CoV) from uninfected controls (U) in diverse human and mouse datasets. **(E)** Time course of CoV infection shows that the ViP host-response signature is slowly induced in very late (48-72 h) in Calu-3 cells infected with SARS-CoV-1. **(F)** ViP signature-based classification of pulmonary infections show that the signature classifies viral lung infection better than bacterial pneumonias from healthy controls with no infections ('*None'*). **(G)** The accuracy (Y axis; ROC AUC) of the signature to classify viral infections differs between RNA viruses and DNA viruses (*X* axis) in *in vitro* system (top). However, they are indistinguishable in *in vivo* system (bottom). See also **Table S2** and **Fig S1**. **(H)** The signature captures host response to CoV infection in human primary lung alveolar epithelial cells (AE) and dendritic cells (DC) better than in Fibroblasts (FI) and Endothelial (ME) cells. The accuracy of classification (ROC-AUC) strongly correlates with ACE2 expression in these cells. **(I)** Classification of macrophage polarization states 'reactive' (M1 polarized), unstimulated M0 and tolerant M2-like samples using the 166-gene ViP signature across diverse datasets.

Figure 4. Identification of a 'severe ViP' signature.

(A) Heatmap of the 166 genes on a dataset (GSE101702) annotated with varying severity of infection (healthy controls, 52; mild, 63; severe, 44). Genes are ranked based on their strength of

association with severity (T-test between mild and severe). Genes that are involved in cytokine signaling in the immune system are highlighted on the left. Heatmap of top 20 selected genes ('severe ViP' signature) is shown on the right. **(B)** Bar and violin plots display sample rank order (i.e., classification) of patient samples and distribution of the 20-gene 'severe ViP' signature in the test dataset (GSE101702). ROC-AUC values of mild and severe cases are shown below the bar plot. **(C)** Reactome pathway analysis of 20 genes. **(D)** Bubble plots of ROC-AUC values (radius of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification based on the 20-gene severe ViP signature (top) and 166-gene ViP signature (bottom) in the test dataset (GSE101702), three more human datasets (H7N9, GSE114466; H1N1, GSE21802; IAV/H3N1 and others, GSE61821) and one mouse dataset (H1N1 Inf A, GSE42641). For each gene signature, ROC-AUC of controls vs Mild and Mild vs Severe are shown in top and bottom rows, respectively. In the mouse dataset (GSE42641) host response to lethal (L) and sublethal (SL) infection with H1N1 virus were assessed in five different lung cell types: Alv Mac, Lymphocytes, Monocytes, Neutrophil, Epithelial cells. Number of controls, mild and severe cases are shown at the top. **(E)** Summary of the 20-gene severe ViP signature and pathway analysis by DAVID GO (https://david.ncifcrf.gov/).

Figure 5. Host immune response in COVID-19

(A) Heatmap of 166 genes in COVID-19 (GSE147507) dataset ranked by genes up-regulated in COVID-19 infected samples. Genes that are involved in cytokine signaling in immune system are highlighted in the left. **(B-G)** Bar and violin plots displaying sample rank order (i.e., classification) and distribution of gene signature scores of COVID-19 (GSE147507) infected (CoV) and uninfected controls (C) in A549 (13 C, 6 CoV; **B, E**), normal human bronchial epithelial cells (NHBE, 7 C, 3 CoV; **C, F**), and patient lung autopsies (2 Normal, 1 CoV; **D, G**) based on 166-gene (**B, C, D**) and 20-gene ViP signatures (**E, F, G**). **(H)** Bubble plots of ROC-AUC values (radius of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification based on the 20 gene severe ViP signature (top) and 166 gene ViP signature (bottom) in multiple independent datasets. **(I)** Bubble plots like panel H showing ROC-AUC of controls vs Mild and Mild vs Severe that are shown in top and bottom rows, respectively for each gene signature in the COVID-19 single-cell datasets (GSE145926).

Dataset is analyzed as a 'pseudo-bulk' of all cells or after selecting individual cell types using marker genes specifically expressed in these cell types.

Figure 6. ViP signatures reveal an interplay between IL15-storm and NK cell dysfunction in fatal COVID-19

(A) Bubble plots of ROC-AUC values (radius of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification based on the 20 gene severe ViP signature (top) and 166 gene ViP signature (bottom) in following datasets. RNASeq data (GSE115203) from PBMCs and sorted NK cells from PBMCs incubated with uninfected A549 cells for 12 hrs compared to infected A549 cells. PBMCs treated with IL15 compared to IL2 (GSE77601). RNASeq analysis of NK cells (GSE89484) treated with GSK-J4 compared to DMSO. Skin tissue in mice (GSE45551) is treated with anti-IL15RB antibody compared to PBS. **(B)** RNASeq data of NK cells isolated from two donors prior to vaccination compared (left) to days 1, 3, and 7 post-TIV vaccination like panel A. RNASeq data of NK enriched and NK depleted PBMCs from healthy donors compared to 30 day post-vaccination like panel A. **(C, D)** Heatmap of 20-gene (panel **C**) and 166-gene (panel **D**) ViP signatures in tissues collected during rapid autopsies on patients who succumbed to COVID-19. Genes are ranked according to the strength of differential expression (T-test) in lung tissue between normal and infected tissue. **(E)** Box plots of IL15 and IL15RA in samples from varying severity of COVID-19. **(F-H)** The level of plasma IL15 cytokines in COVID-19 patients were analyzed after hospital admission. Violin plots in show levels of plasma IL15 in patients stratified by disease acquity (F), by clinical severity (G) and by gender and age (H). See also **Table S4** for patient metadata.

Figure 7. **Lung alveolar cells contribute to the IL15 storm in fatal COVID-19**

(**A**) Normal lung tissue obtained during surgical resection (*left*) or lung tissue obtained during autopsy studies on COVID-19 patients (*right*) were stained for IL15 and IL15RA. Representative images are shown. Mag = 10X. (**B**) Violin plots display the intensity of staining for IL15RA (top) and IL15 (bottom), as determined by IHC profiler. (**C**) Summary of IL15 signaling and the hypothetical role of NK cells in the severity of COVID-19 infections.

Figure 8. Validation of ViP signature-guided therapeutic goals

(A-C) The 166-gene ViP signature-was used to classify liver biopsies from HCV-infected patients treated or not with directly acting anti-viral agents. ROC-AUC values are shown below each bar plot unless otherwise stated. **(D)** 166-gene ViP signature-based classification of blood samples from HIV-infected patients treated with anti-retroviral therapy (ART). **(E)** 166-gene ViP signature-based classification of crisis and convalescence in PBMCs from patients with Ebola infection. **(F)** The effect of inhibiting Topoisomerase 1 (iTop) in a cultured cell line model infected in vitro with Ebola for the development of anti-Ebola therapeutics. **(G)** 166-gene ViP signature-based classification of human cortical neural progenitor cells infected *in vitro* with Zika virus (*Left*) and the infected cells treated with two investigational drugs (*Right*; Rx 1 and 2) during screening assays. **(H)** Schematic showing the experimental design for validating the ViP signatures as useful tools to assess therapeutic efficacy. *Uninf*, uninfected; *Den3 and Anti-CoV-2* indicate SARS-CoV-2 challenged groups that received either a control mAb or the clone CC12.2 of anti-CoV-2 IgG, respectively. (**I**) Bar (*top*) and violin (*bottom*) plots display the 166- and 20-gene ViP signatures in the uninfected and the SARS-CoV-2 challenged groups, treated with control or anti-CoV-2 IgG. (**J-L**) Lungs harvested from the 3 groups of hamsters were analyzed by H&E and IHC. Representative images are shown in J. Mag = 10X. Bar graphs in K display the abundance of cellularity and infiltrates in the lungs of the 3 groups, as determined by ImageJ. Violin plots in L display the intensity of staining for IL15RA (top) and IL15 (bottom), as determined by IHC profiler.

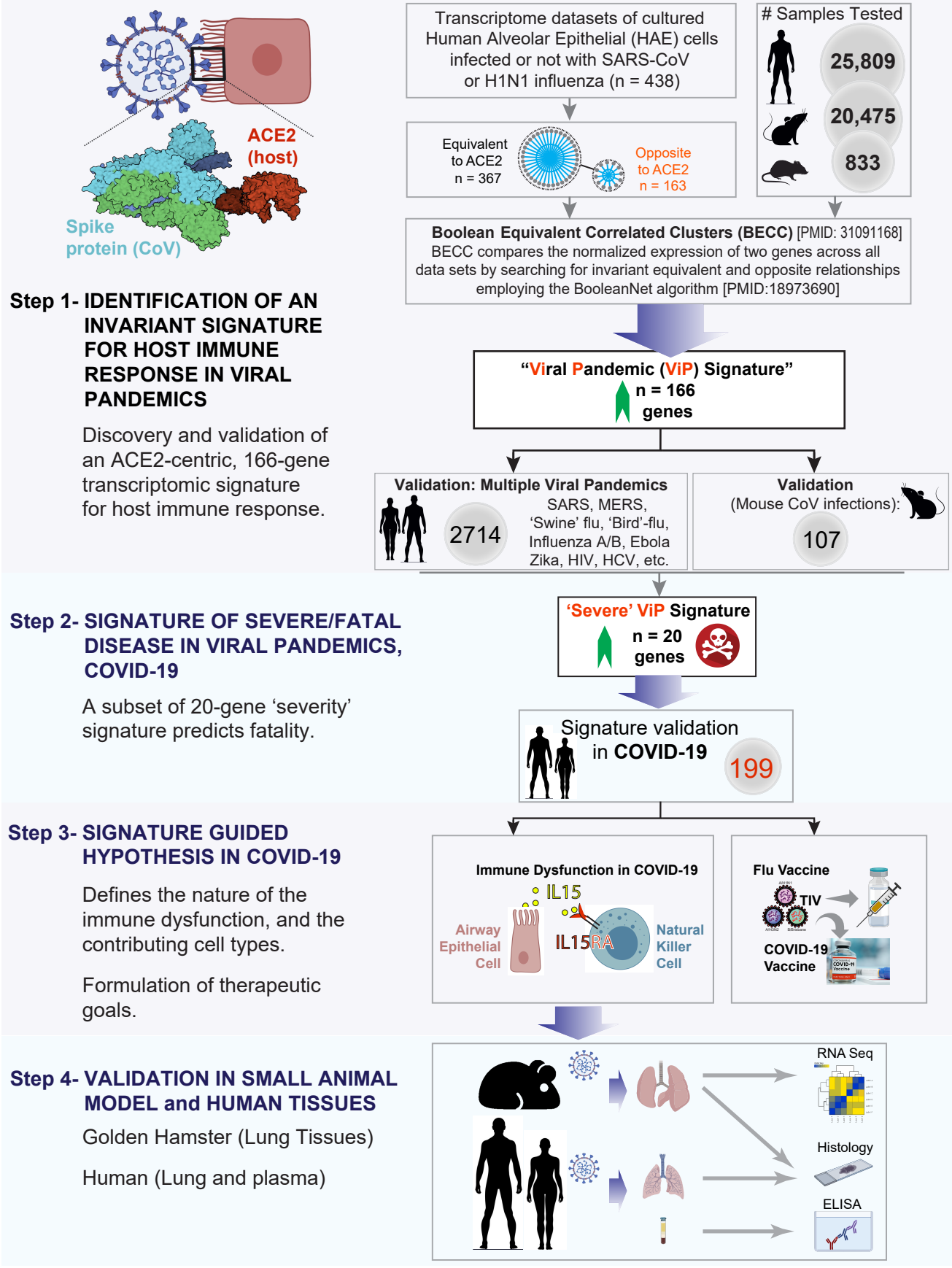
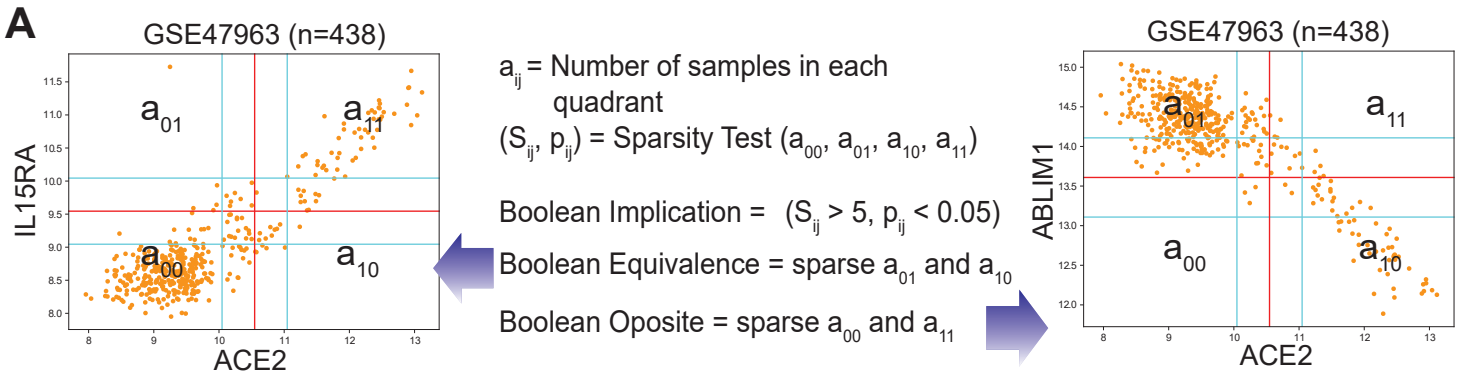
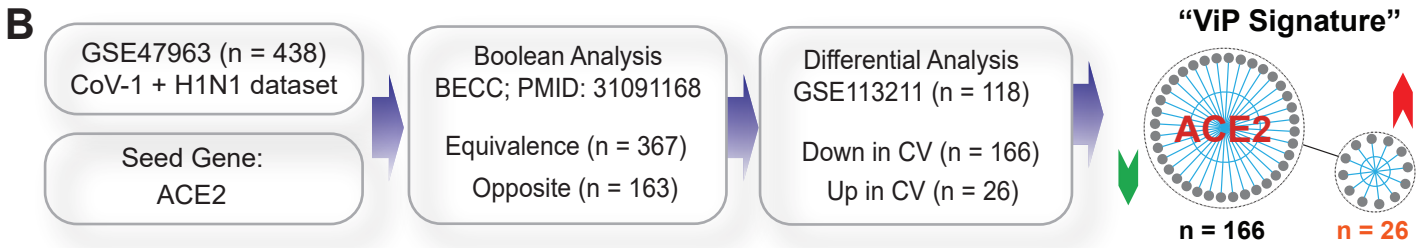


Figure 1





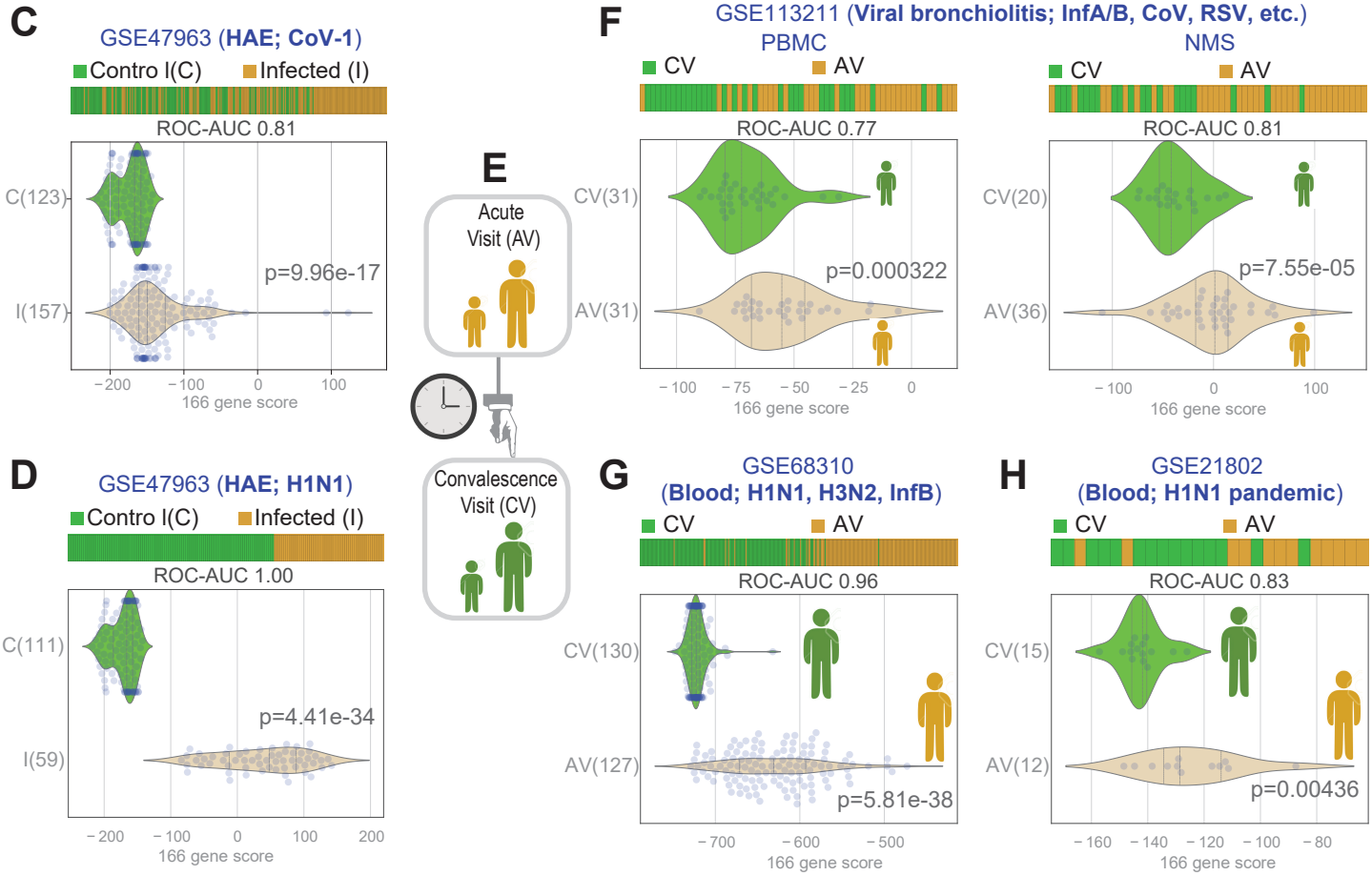


Figure 2

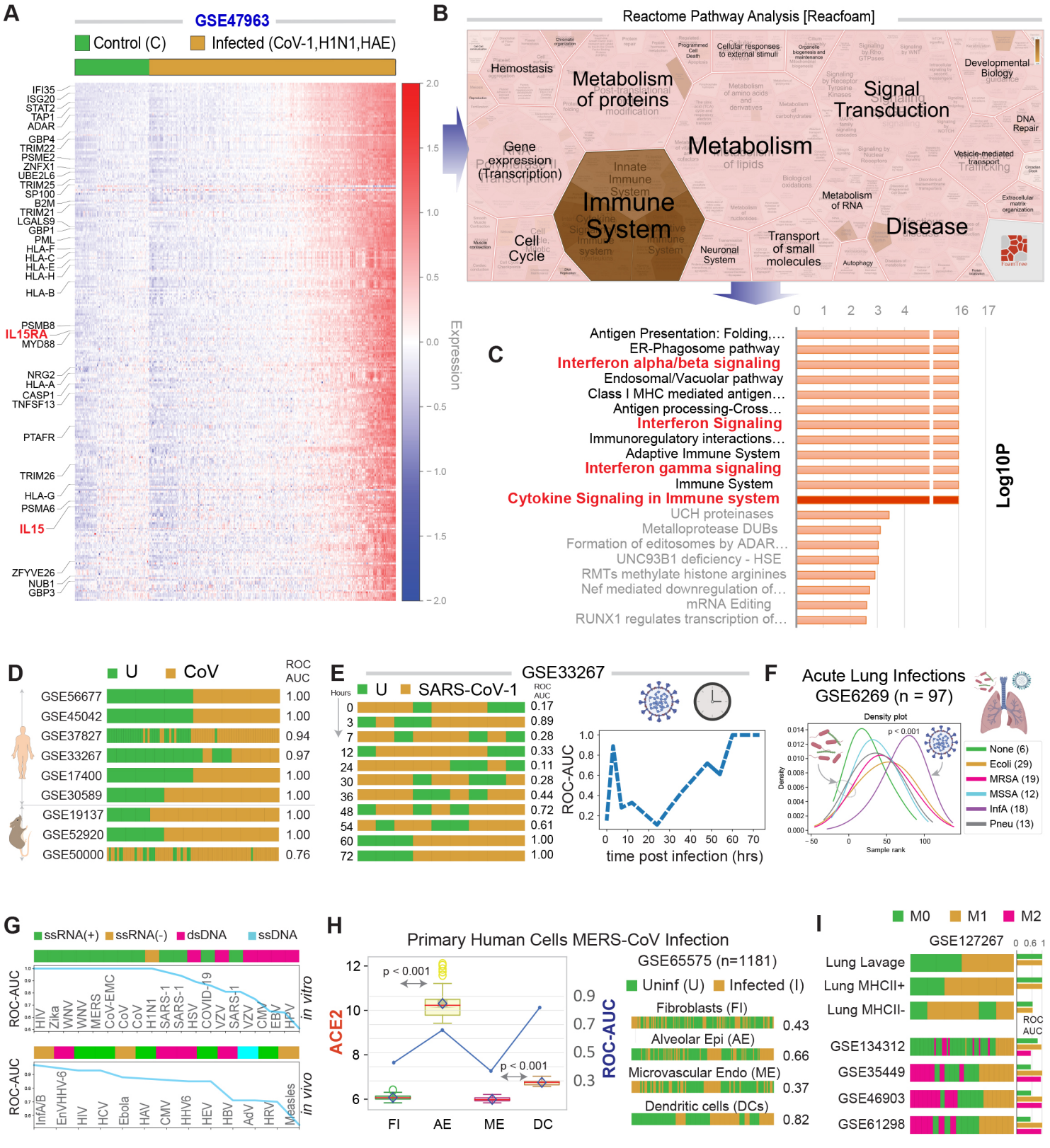


Figure 3

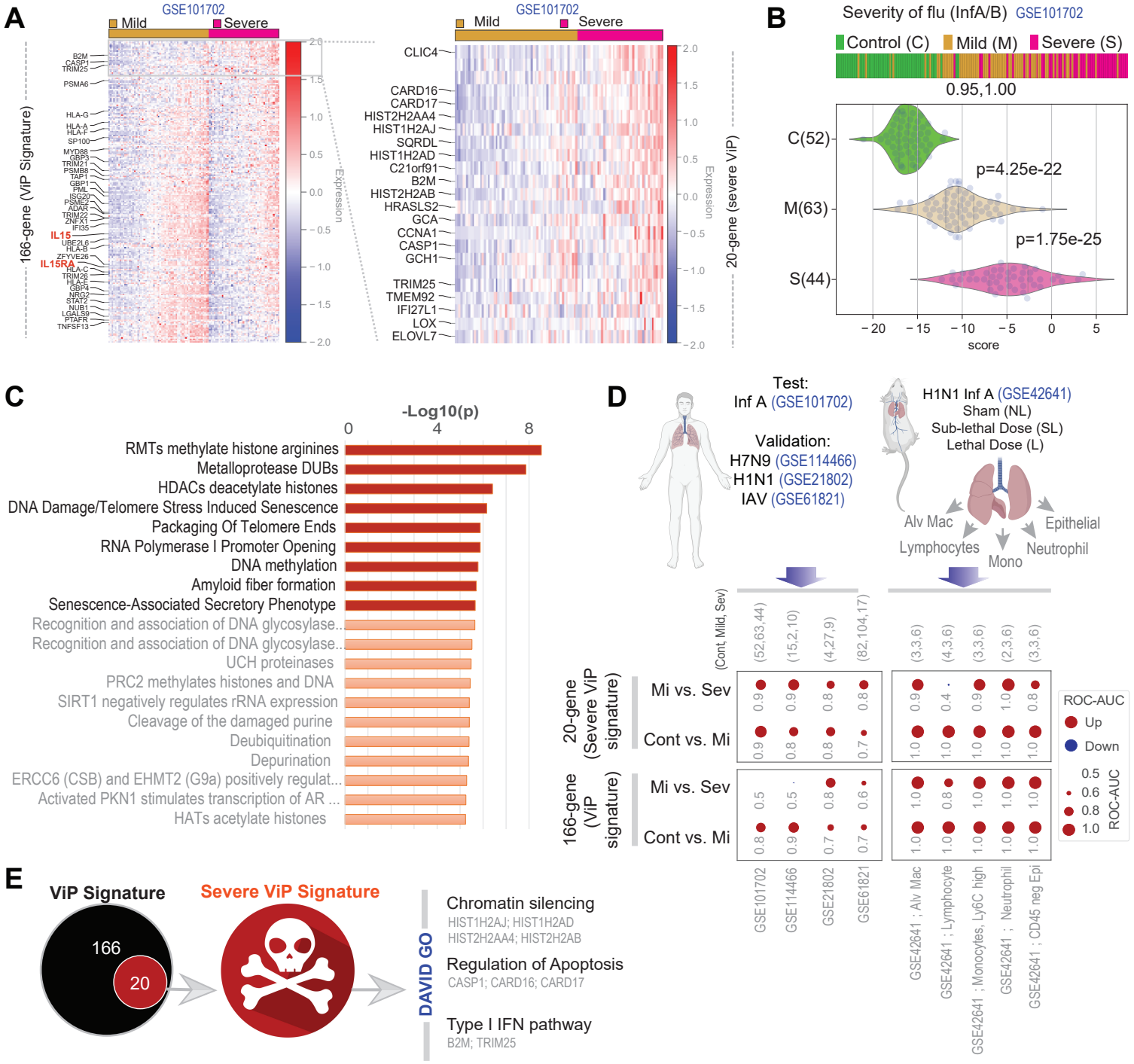


Figure 4

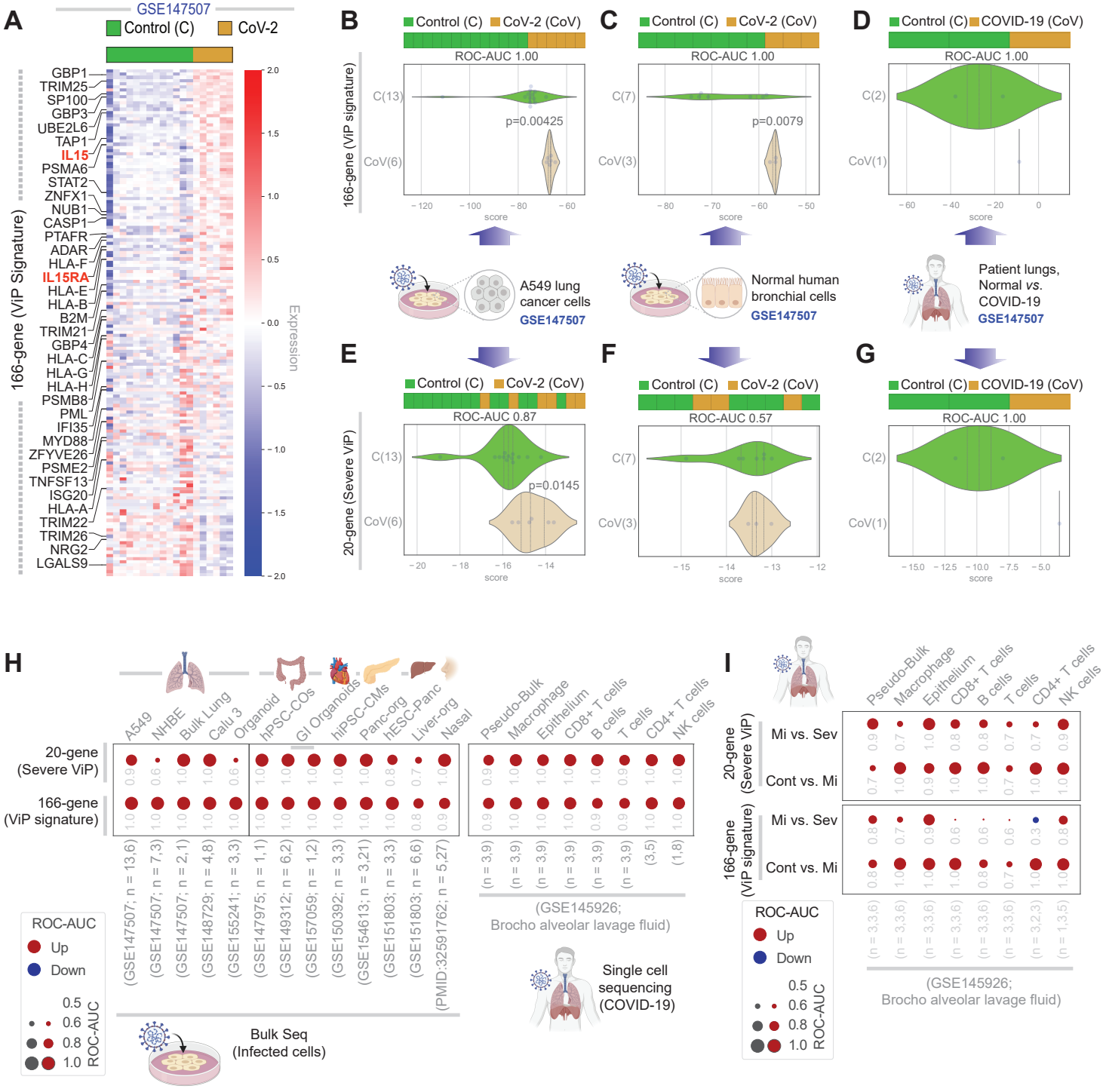


Figure 5

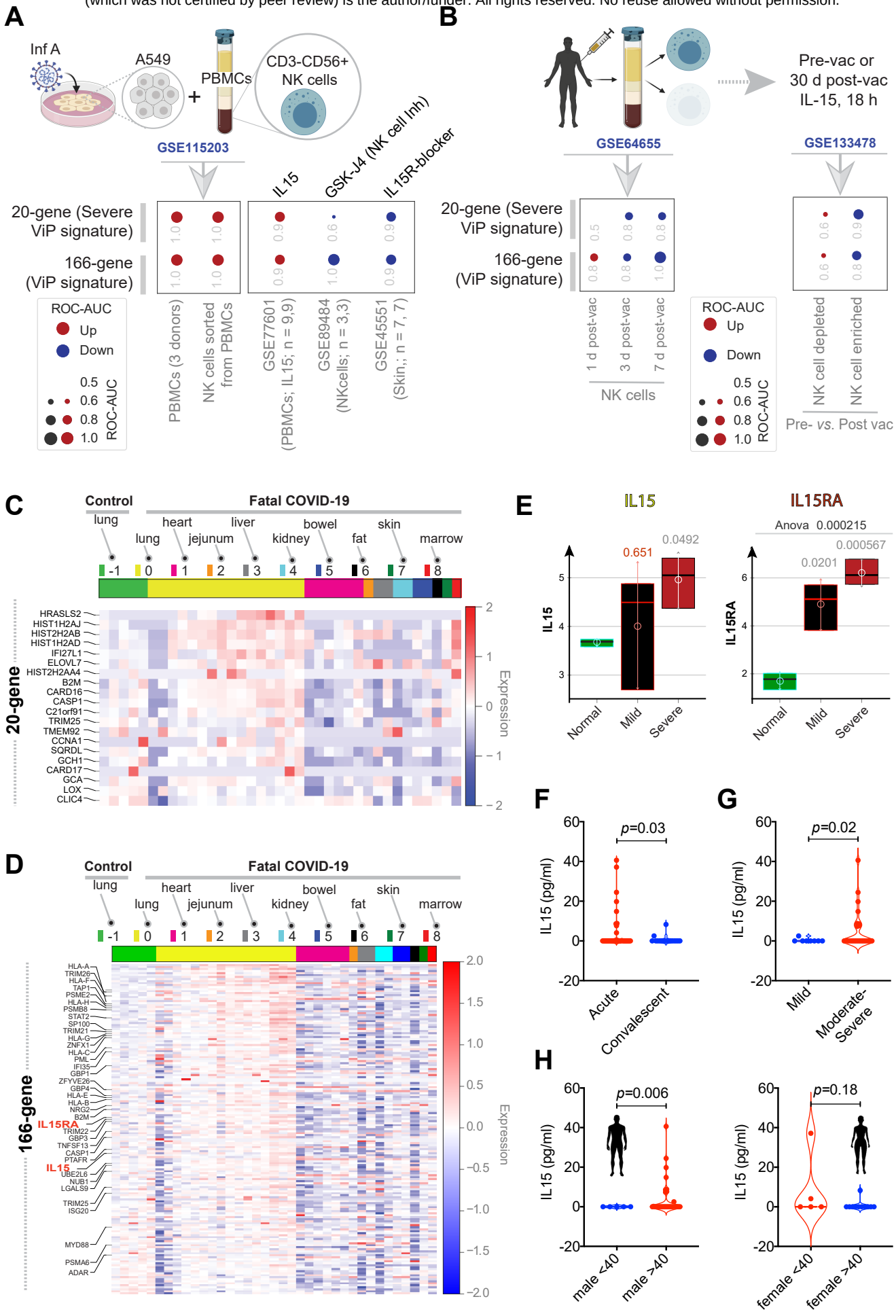


Figure 6

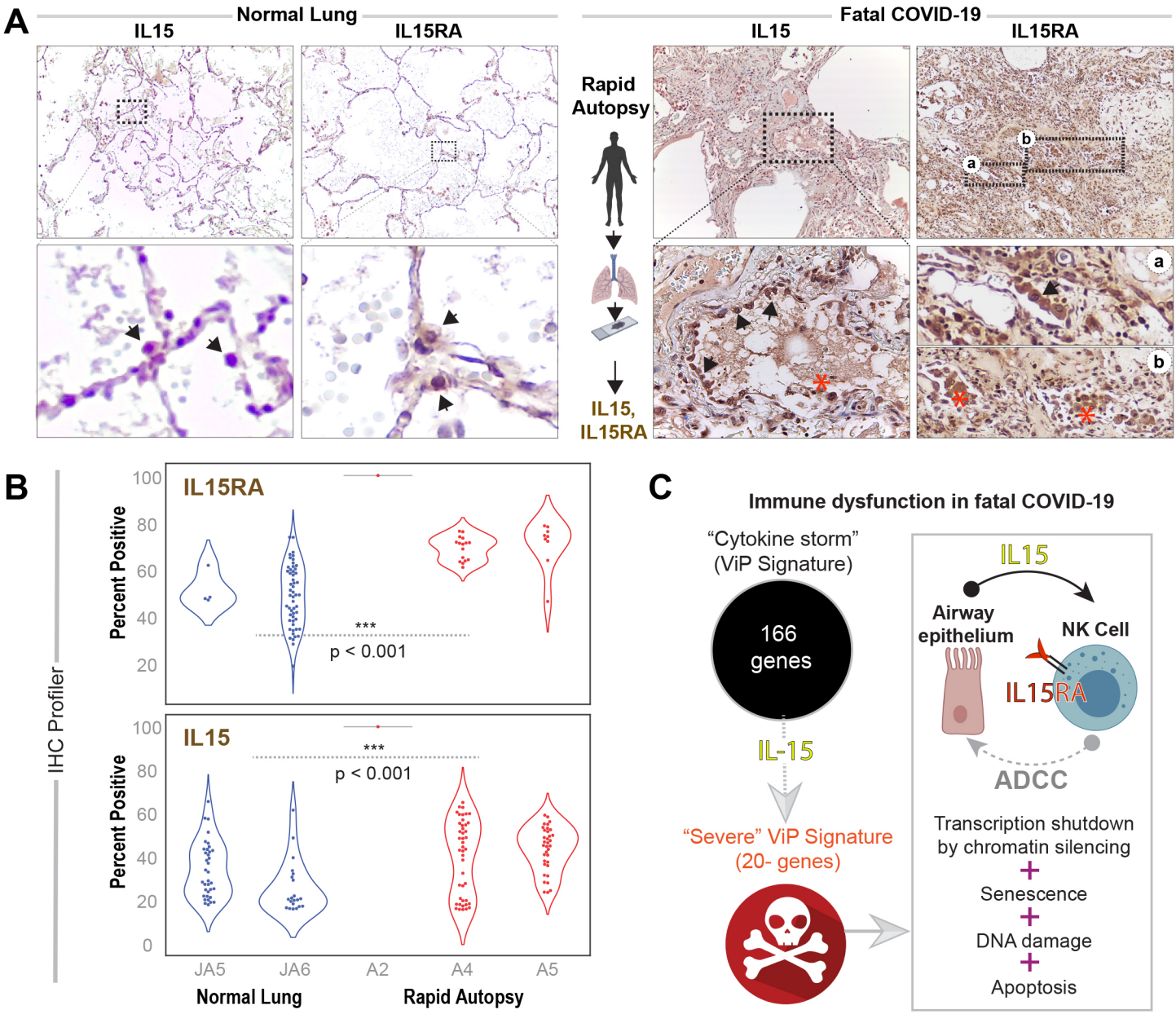
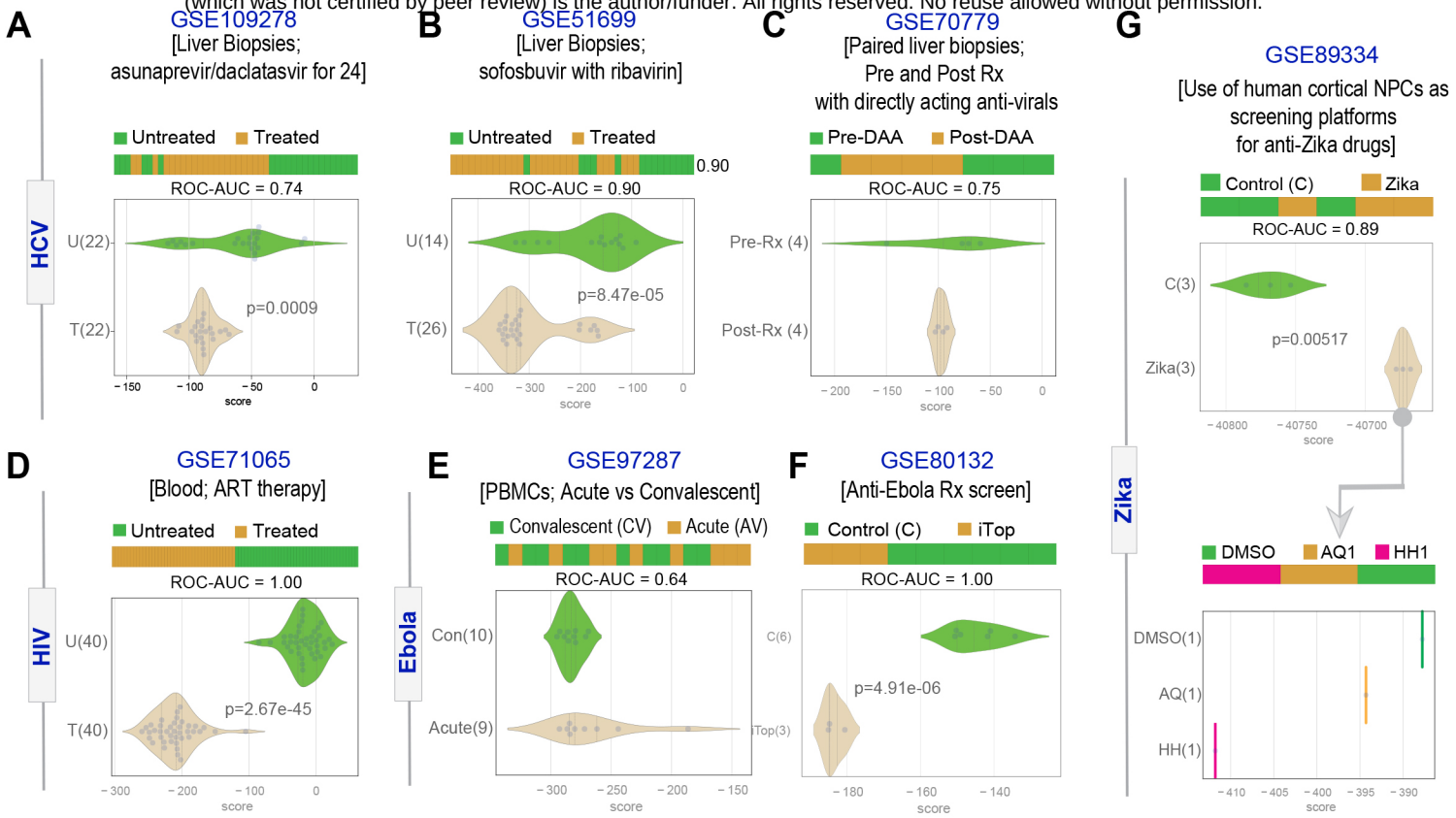
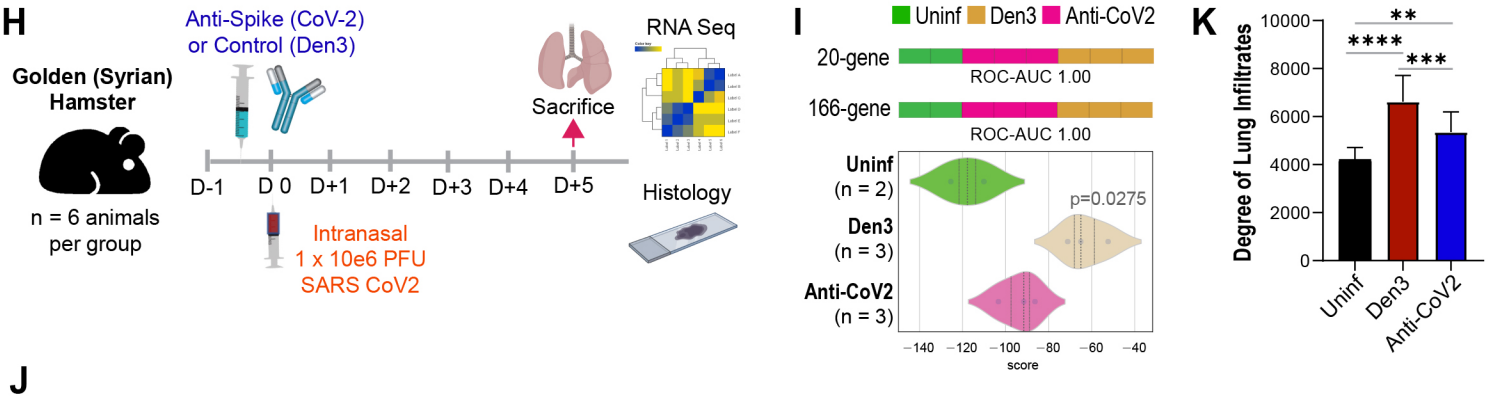


Figure 7





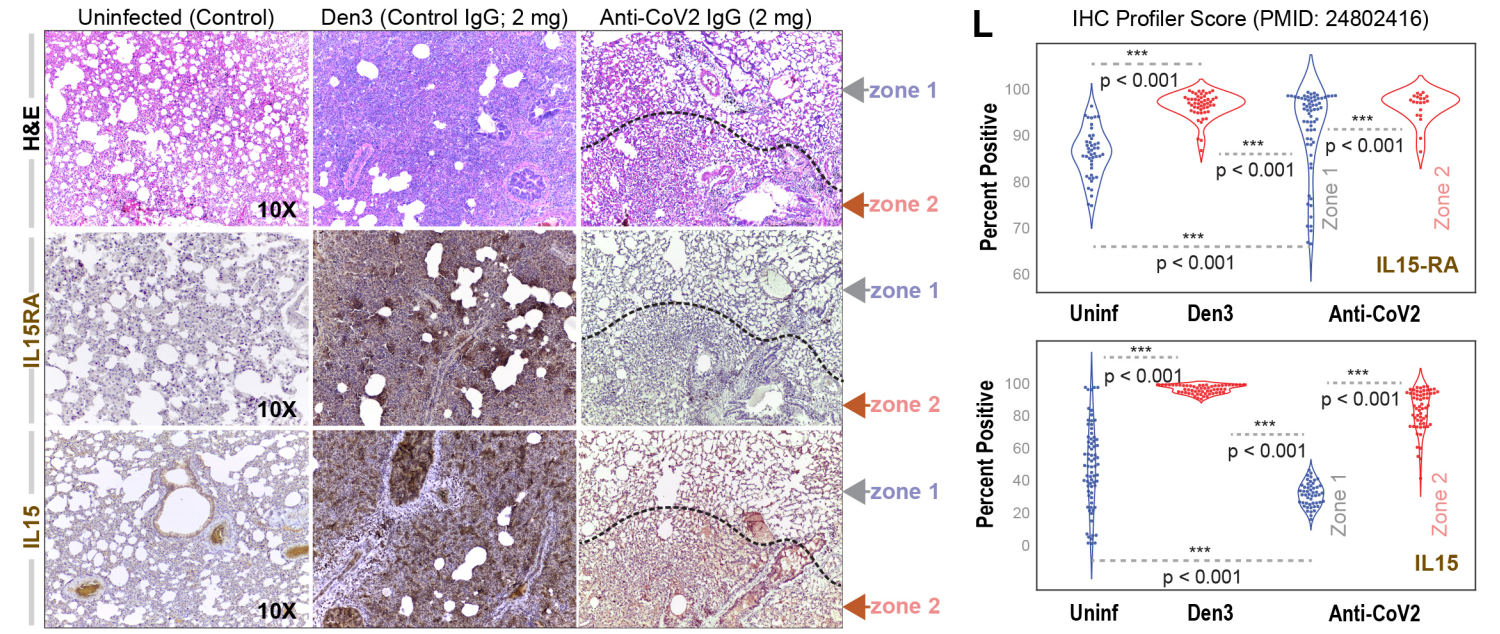


Figure 8