# Meta-Analysis Identifies Type I Interferon Response as Top Pathway Associated with SARS Infection

**Amber Park** *ampark10@gmail.com*
*Science Department*
*Davenport University*
*Grand Rapids, 49512, USA*

**Marius Nwobi** *nwobimar@msu.edu*
*Biochemistry Department*
*Michigan State University*
*East Lansing, 48824, USA*

**Laura K. Harris** *oesterei@msu.edu*
*Institute for Cyber Enabled Research*
*Michigan State University*
*East Lansing, 48824, USA*

**Abstract**

**Background**: Severe Acute Respiratory Syndrome (SARS) corona virus (CoV) infections are a serious public health threat because of their pandemic-causing potential. This work examines pathway signatures derived from mRNA expression data as a measure of differential pathway activity between SARS and mock infection using a meta-analysis approach to predict pathways associated with SARS infection that may have potential as therapeutic targets to preclude or overcome SARS infections. This work applied a GSEA-based, meta-analysis approach for analyzing pathway signatures from gene expression data to determine if such an approach would overcome FET limitations and identify more pathways associated with SARS infections than observed in our previous work using gene signatures.

**Methods**: This work defines 37 pathway signatures, each a ranked list of pathway activity changes associated with a specific SARS infection. SARS infections include seven SARS-CoV1 strains with established mutations that vary virulence (infectious clone SARS (icSARS), Urbani, MA15, ΔORF6, Bat-SRBD, ΔNSP16, and ExoNI), MERS-CoV, and SARS-CoV2 in human lung cultures and/or mouse lung samples. To compare across signatures, positive and negative icSARS pathway panels are defined from shared leading-edge pathways identified by Gene Set Enrichment Analysis (GSEA) between two icSARSvsmock signatures, both from human cultures. GSEA then assesses enrichment in all 37 signatures and identifies leading-edge icSARS panel pathways for each analysis. A meta-analysis across identified leading-edge pathways reveals commonalities which are ranked by Stouffer's method for combining p-values.

**Results**: Significant enrichment (GSEA $p<0.001$) is observed between the two icSARSvsmock signatures used to define positive (195 pathways) and negative (173 pathways) icSARS panels. Consistent, non-random (null distribution $p<0.01$), significant enrichment of the positive icSARS pathway panel in all pathway signatures is observed but significant enrichment is inconsistent for the negative icSARS panel. After meta-analysis, 11 pathways are found in all GSEA-identified leading-edges from the positive icSARS panel. Identified pathways are involved in the immune system with response to type I interferon ranked highest.

**Conclusion**: This GSEA-based meta-analysis approach identifies pathways with and without reported associations with SARS infections, highlighting this approach's predictability and usefulness in identifying pathway activity changes.

**Keywords:** Pathway Activity, Meta-analysis, Gene Set Enrichment Analysis, SARS, Interferon.

## 1. INTRODUCTION

A group of positive-sense RNA viruses that infect humans and a variety of animal species, known as human β-coronaviruses (CoV), typically cause mild upper respiratory distress referred to as the common cold. Many professionals viewed this viral group as non-lethal and no great concern for several decades [1]. This view changed in 2002 with a novel CoV that caused a life-threatening disease called severe acute respiratory syndrome (SARS). The SARS-CoV1 epidemic infected over 8,400 people and had over a 9% mortality rate. In 2012, another CoV emerged causing Middle East Respiratory Syndrome (MERS), which had an over 30% mortality rate [2]. However, it has been noted that MERS-CoV is part of CoV lineage C which is phylogenetically different from other human CoV including SARS-CoV1 and utilizes dipeptidyl peptidase 4 (DPP4) rather than angiotensin-converting enzyme 2 (ACE2) for cellular entry [3]. Thankfully both SARS-CoV1 and MERS-CoV epidemics were contained by countries who took with advanced technological collaborations and regulated animal husbandry practices [4]. SARS-CoV2, the causative agent of CoV disease 2019 (COVID-19), is responsible for over 3.7 million deaths across the globe by early June 2021 and continues to cause a raging pandemic [5]. SARS-CoV2 has an 80% genome similarity to SARS-CoV1 and a 50% similarity to MERS-CoV [6, 7]. New variants of SARS-CoV2 are reported regularly, and some variants like B.1.1.7 are reported to have increased infectiousness compared to older variants, forcing scientists to consider outbreaks of a future SARS-CoV3 strain [8-11].

Successfully treating SARS-infected patients is crucial to improving patient conditions and restoring communities. Unfortunately, there were limited therapies for SARS-CoV1 and MERS-CoV infections available. Corticosteroids, RNA-dependent RNA polymerase inhibitors, and nucleoside antimetabolite drugs like ribavirin were the main treatments against SARS-CoV1; however, these specific drugs have displayed poor efficacy or questionable effectiveness [12-14]. Remdesivir is a nucleotide analogue that inhibits viral RNA synthesis in SARS by interfering with RNA polymerase and evading viral exoribonuclease [15]. Remdesivir has shown inconsistent results with some studies not showing ay clinical benefits or difference in 5- or 10- day regiments with a potential stronger effect in severe versus moderate patients. Treatments applied for SARS-CoV1, including ribavirin and remdesivir, were utilized on MERS-CoV with little investment into developing new therapeutic options. Despite existing drugs being repurposed to target MERS-CoV replication in vitro, there were limited treatment options available for patients at the time. The urgency to develop therapies against SARS increased substantially after the appearance of SARS-CoV2 as demonstrated through the execution of over 284 clinical trials with the goal to examine the efficacy and safety of repurposed and novel drugs to treat SARS infections [4]. A few treatment options include repurposing past drugs that displayed high potential efficacy against SARS-CoV1 and MERS-CoV infections [16]. Specific drugs such as remdesivir, favipiravir, ivermectin, lopinavir/ritonavir, arbidol, and others that inhibit membrane fusion, viral replicases, and human and viral proteases also have been under consideration as ways to manage and treat SARS-CoV2 [17, 18]. Newer therapies are incorporating immunotherapy with some therapeutics specifically targeting the cytokine storm consistently reported in COVID 19 patients. For example, interferon (IFN) and interleukin (IL)-6 receptor antagonist therapies are beginning to emerge as strong clinical options [12-14, 18-21]. Regarding preventative measures, several novel and traditional vaccinations have been approved for administration with many other vaccine candidates currently in clinical trials [18, 20]. Despite current vaccinations having high efficacy, no definite treatment option is available to consistently treat SARS-CoV2 infections as death tolls from the virus continue to increase [19]. New therapeutic strategies are vital to successful treatment of current and future SARS infections.

Analyzing pathway activity changes associated with SARS infection can identify new targets for the development of new therapies to combat COVID-19 and future SARS outbreaks [22]. Several studies have examined pathway activity changes in SARS infections by analyzing gene

expression data [3, 23-30]. Pathway activity changes usually are identified individually using a hypergeometric test variation, such as Fisher's Exact Test (FET), or Gene Set Enrichment Analysis (GSEA) [22]. FET calculates enrichment of a pathway (*i.e.*, established list of genes associated with a certain cellular product or change) in a researcher-defined gene subset, such as differentially expressed genes identified via statistical cut-offs such as fold change>2 and/or T-test p-value<0.05. Pathway enrichment analysis using hypergeometric test variation on differentially expressed genes identified from SARS data has revealed increased immune response, chemokine, and cytokine pathway activities, particularly IL-6 and IL-8 pathways, in SARS-CoV1 infected human lung cultures and mouse lung samples [26], increased inflammatory, coagulation, and apoptotic pathway activity and decreased IFN-I signaling pathway activity in SARS-CoV2 infected human lung cell cultures and COVID-19 patients [23-25, 27, 28]. While hypergeometric tests are commonly used, GSEA has become a preferred method to calculate pathway enrichment because it considers all genes in an expression dataset rather than only differentially expressed genes meeting an established statistical cut off [22, 31]. GSEA accomplishes this through examination of gene signatures, which are ranked lists of expression dataset genes based on differential gene expression [31]. Enrichment of individual innate immune response pathways taken from public knowledgebases (*e.g.*, Gene Ontology (GO) [32], Kyoto Encyclopedia of Genes and Genomes (KEGG) [33], and/or MSigDB [34]) was found consistently in the positive tails of SARS gene signatures derived from blood, lung, and airway samples of COVID-19 patients [29, 30]. Comparative transcriptome analysis across SARS strains identified similarities and differences in which pathways were significant individually; for example, finding increased phagosome pathway in MERS-CoV and not SARS-CoV1 or SARS-CoV2 [3]. While these pathway enrichment approaches have been successful in identifying cellular mechanisms driving SARS infections, the approaches are limited because they consider each pathway individually. New computational approaches that consider activity changes of all pathways simultaneously may offer a new perspective into how SARS alters a cell, potentially leading to identification of new therapeutic strategies.

In a prior work, we compared gene signatures across SARS strains using a GSEA-based meta-analysis approach [35]. This approach successfully identified differentially expressed genes with and without prior association to SARS infections with five IFN-induced genes consistently identified across 37 SARS gene signatures. Our prior study expanded upon previous work by overcoming limitations of single-gene analysis (e.g., T-test) through using GSEA to identify genes. However, pathway analysis on identified genes in that study was limited by use of a FET-variant to identify pathway activity changes associated with SARS infection. To overcome that FET limitation, in this work we performed a GSEA-based meta-analysis that used pathway signatures to identify pathways associated with SARS infections from mRNA expression datasets. By comparing pathway signatures using GSEA we can consider all pathways simultaneously and acquire a more complete understanding of cellular changes that occur during SARS infection. Further, this improved understanding could reveal molecular insights for the development of therapeutic strategies against COVID-19 and future SARS outbreaks though we acknowledge that this work is entirely computational so further experimental and clinical work will be needed to properly validate and implement our findings.

## 2. METHODS

### 2.1 mRNA Expression Resources and Previously Defined Gene Signatures

In our recent gene expression study [35], we selected 17 datasets from Gene Expression Omnibus (GEO) [36] that contained SARS or mock infected human lung cell cultures or mouse lung samples collected at the 48hr time point post-infection. To briefly describe these datasets, 11 datasets contained SARS-CoV1 cultures and samples across seven strains with varying virulence levels and infectious doses [26, 37-39]. These strains included a(n) 1) wild-type highly virulent strain Urbani, 2) infectious clone of SARS-CoV1 Urbani (icSARS), 3) MA15, a mouse adapted SARS-CoV1 strain, 4) attenuated icSARS with a genetic deletion causing lack of expression of IFN antagonist protein ORF6, 5) SARS-CoV1 like virus isolated from bats that was synthetically modified to contain the spike receptor binding domain (SRBD) from wild type Urbani for infection

of human cells (BatSRBD), 6) icSARS with a genetic modification of NSP16 (deltaNSP16) that encodes a non-structural 2'O methyltransferase whose disruption increases sensitivity to type I and III IFN responses, and 7) SARS-CoV1 with a genetic modification of ExoNI, which has no formal description in public knowledgebases. Three datasets contained MERS-CoV cultures and samples with varying infectious doses, and three datasets contained SARS-CoV2 cultures, presumably with the same infectious dose [40]. These datasets were selected so all available GEO datasets containing SARS and mock samples collected at 48hrs post infection were included in our meta-analysis.

Our prior study used these 17 datasets to define 37 gene signatures which we used as previously described to define pathway signatures in this study [35]. To briefly describe these gene signatures, there were 29 SARS-CoV1 gene signatures which included 17 gene signatures from human lung cultures and 12 signatures from mouse lung samples. Specifically, we used seven icSARSvsmock signatures (six in human cultures, one in mouse samples), one UrbaniSARSvsmock signature with the Urbani strain in human cultures, eight MA15vsmock signatures doses in mouse samples representing varying inoculation, five ORF6vsmock signatures (four in human cultures, one in mouse samples), five BATvsmock signatures (four in human cultures, one in mouse samples), two NSP16vsmock signatures (one in human cultures, one in mouse samples), and one ExoNIvsmock signature from human cultures. For MERS-CoV datasets, we utilized three datasets to define five gene signatures that included two gene signatures from human lung cultures and three signatures from mouse lung samples. For SARS-CoV2 datasets, we used three datasets to define three gene signatures, all from human lung cultures. We further adopted the same research design, data collection, and data analysis as our prior study to inductively determine if a pathway signature meta-analysis approach would include and expand upon our previous gene expression based IFN findings.

### 2.2 Defining Pathway Signatures using Gene Set Enrichment Analysis

To measure pathway activity changes from a gene signature, we converted each gene signature into a pathway signature using Gene Set Enrichment Analysis (GSEA). GSEA is a statistical method that estimates enrichment between a query gene set (*i.e.*, unranked list of genes) and a reference gene signature [31]. GSEA uses the statistical metric used to rank genes in the reference signature to calculate a running summation enrichment score where hits (*i.e.*, matches between query set and reference signature) increase the enrichment score proportional to the ranking statistical metric and a miss (*i.e.*, non-matches between query set and reference signature) decreases the enrichment score. From this, GSEA determines a maximum enrichment score for the specific query set and reference signature. Leading-edge genes contribute to reaching the maximum enrichment score, indicating leading-edge genes are associated with cellular response to a specific SARS infection. Further, GSEA calculates a normalized enrichment score (NES) from 1000 permutations of the reference signature to estimate the significance of enrichment between a specific query set and reference signature. By using the gene lists from an established pathway knowledgebase, pathway activity changes can be estimated by NES and ranked accordingly. Therefore, we examined available gene sets in the Molecular Signatures Database (MSigDB) version 7.2, which is curated by the Broad Institute, and found the GO BP collection had the most with 7573 gene sets. We used these GO BP gene lists as individual query sets (Supplemental Material STable 1) for GSEA against each gene signature (reference) to generate NES. Pathway signatures (*i.e.*, a list of pathways ranked from high to low differential activity between two groups, such as SARS and mock infected lung cells) were generated by ranking pathways by NES while removing pathways that were under-represented in the gene signature (*i.e.*, pathway with less than 15 genes in the gene signature). Pathways that were up-regulated in SARS compared to mock infected samples (*e.g.*, positive NES, red color) fall within the positive tail of the pathway signature while down-regulated pathways (*e.g.*, negative NES, blue color) fall in the negative tail (Figure 1). Pathways with no substantial change in activity between SARS and mock infected samples (*e.g.*, NES around 0) were located toward the middle of the pathway signature. Therefore, pathways that fall within the tails of a pathway signature likely changed in response to a specific SARS infection. Pathway signature information was included in Tables 1 and 2. The same mock samples were used across signatures in same

dataset. Values in Length column were total number of pathways in the signature and Cross column reflected where midpoint was in the signature. This work used the javaGSEA Desktop Application release 3.0 version of GSEA available from Broad Institute.

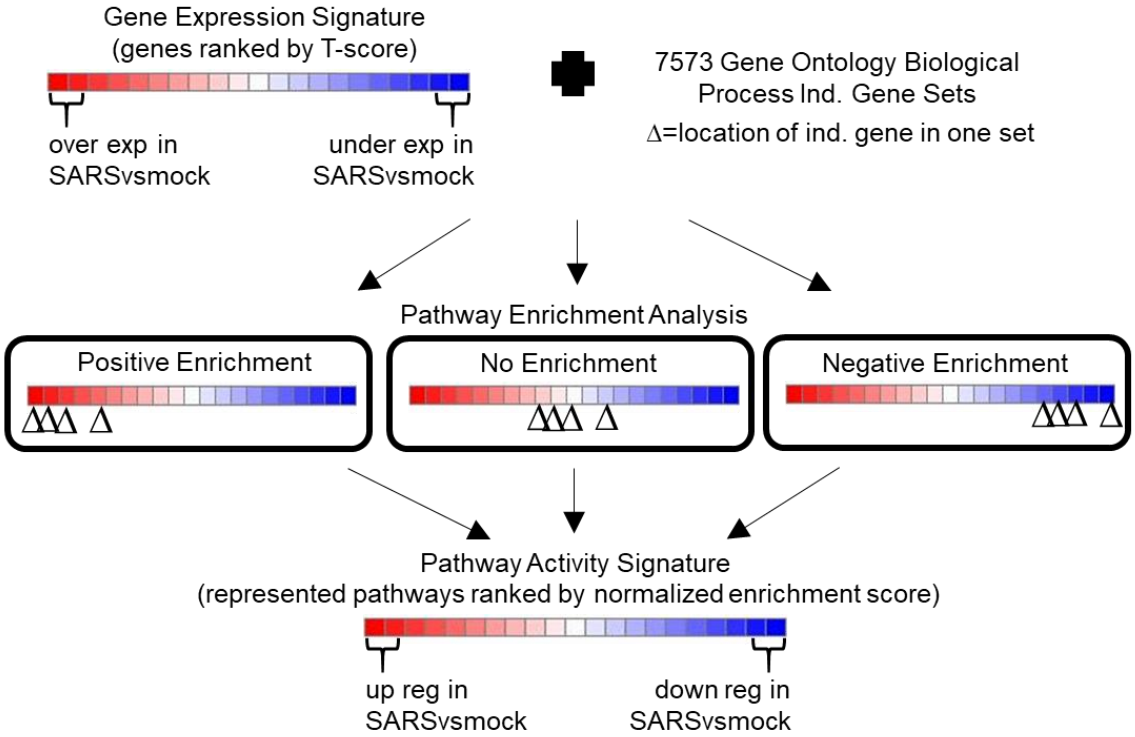


**FIGURE 1:** Schematic Definition of a Pathway Signature.

| Dataset | Group 1 (number of samples) | Group 2 (number of samples) | Pathway signature | Use | Max NES | Minimum NES | Length | Cross |
|---|---|---|---|---|---|---|---|---|
| GSE47960 | icSARS (4) | mock (3) | icSARSvsmock | Identification | 2.48 | -2.55 | 3871 | 2141 |
| | dORF6 (4) | mock (3) | ORF6vsmock | Comparison | 2.47 | -2.69 | 3871 | 2760 |
| | BatSRBD (3) | mock (3) | BATSRBDvsmock | Comparison | 2.05 | -2.14 | 3871 | 1668 |
| GSE47961 | icSARS (4) | mock (3) | icSARSvsmock | Identification | 2.81 | -2.74 | 3893 | 1541 |
| | dORF6 (4) | mock (3) | ORF6vsmock | Comparison | 2.56 | -2.94 | 3893 | 2292 |
| | BatSRBD (4) | mock (3) | BATSRBDvsmock | Comparison | 2.43 | -2.45 | 3893 | 1588 |
| GSE47962 | icSARS (3) | mock (3) | icSARSvsmock | Verification | 2.18 | -2.68 | 3861 | 2462 |
| | dORF6 (3) | mock (3) | ORF6vsmock | Comparison | 2.07 | -3.00 | 3861 | 2515 |
| | BatSRBD (3) | mock (3) | BATSRBDvsmock | Comparison | 2.24 | -2.84 | 3861 | 2607 |
| GSE37827 | icSARS (3) | mock (3) | icSARSvsmock | Verification | 3.00 | -2.41 | 4112 | 2313 |
| | BatSRBD (3) | mock (3) | BATSRBDvsmock | Comparison | 3.02 | -2.20 | 4112 | 2597 |
| GSE48142 | icSARS (3) | mock (3) | icSARSvsmock | Verification | 3.31 | -2.49 | 3880 | 2140 |
| | ExoNI (3) | mock (3) | ExoNIvsmock | Comparison | 3.29 | -2.54 | 3880 | 1978 |
| | dNSP16 (3) | mock (3) | NSP16vsmock | Comparison | 3.26 | -2.21 | 3880 | 2085 |
| GSE33267 | icSARS (3) | mock (3) | icSARSvsmock | Verification | 3.31 | -2.84 | 4109 | 2642 |
| | dORF6 (3) | mock (3) | ORF6vsmock | Comparison | 2.57 | -3.07 | 4109 | 2974 |
| GSE17400 | Urbani (3) | mock (3) | Urbanivsmock | Comparison | 3.05 | -2.35 | 4129 | 1360 |
| GSE50000 | icSARS (5) | mock (4) | icSARSvsmock | Comparison | 2.57 | -2.15 | 3914 | 1541 |
| | $MA10^5$ (4) | mock (4) | $MA10^5$vsmock | Comparison | 3.25 | -2.70 | 3914 | 2025 |
| | $MA10^4$ (4) | mock (4) | $MA10^4$vsmock | Comparison | 2.71 | -2.17 | 3914 | 2093 |
| | BatSRBD (5) | mock (4) | BATSRBDvsmock | Comparison | 2.34 | -1.92 | 3914 | 1875 |
| GSE33266 | $MA10^5$ (5) | mock (3) | $MA10^5$vsmock | Comparison | 2.50 | -2.07 | 3745 | 3140 |
| | $MA10^4$ (5) | mock (3) | $MA10^4$vsmock | Comparison | 2.81 | -2.47 | 3745 | 3276 |
| | $MA10^3$ (5) | mock (3) | $MA10^3$vsmock | Comparison | 2.73 | -1.97 | 3745 | 3407 |
| | $MA10^2$ (5) | mock (3) | $MA10^2$vsmock | Comparison | 2.74 | -2.12 | 3745 | 3149 |
| GSE49262 | $MA10^5$ (3) | mock (3) | $MA10^5$vsmock | Comparison | 2.73 | -2.13 | 3914 | 2599 |
| | dORF6 (3) | mock (3) | ORF6vsmock | Comparison | 2.80 | -2.03 | 3914 | 2675 |
| GSE49263 | $MA10^5$ (4) | mock (3) | $MA10^5$vsmock | Comparison | 2.69 | -2.11 | 3914 | 2572 |
| | dNSP16 (4) | mock (3) | NSP16vsmock | Comparison | 2.74 | -2.22 | 3914 | 2550 |

**TABLE 1:** SARS-CoV1 Pathway Signatures Defined in this Study.

| Dataset | Group 1 (number of samples) | Group 2 (number of samples) | Pathway signature | Use | Max NES | Minimum NES | Length | Cross |
|---|---|---|---|---|---|---|---|---|
| GSE81909 | icMERS (5) | mock (5) | icMERSvsmock | Comparison | 2.96 | -1.80 | 4124 | 3048 |
| GSE100504 | icMERS (5) | mock (5) | icMERSvsmock | Comparison | 2.81 | -2.75 | 4125 | 2735 |
| GSE108594 | MERS (4) | mock (4) | $MERS10^{4}$vsmock | Comparison | 2.60 | -2.09 | 3923 | 2290 |
| | MERS (4) | mock (4) | $MERS10^{5}$vsmock | Comparison | 2.54 | -2.06 | 3923 | 2153 |
| | MERS (4) | mock (4) | $MERS10^{6}$vsmock | Comparison | 2.57 | -2.15 | 3923 | 1804 |
| GSE152586 | SARS2 (3) | mock (3) | SARS2vsmock | Comparison | 2.57 | -2.56 | 4180 | 2218 |
| GSE160435 | SARS2 (5) | mock (5) | SARS2vsmock | Comparison | 2.65 | -2.05 | 4210 | 2992 |
| GSE155518 | SARS2 (3) | mock (3) | SARS2vsmock | Comparison | 2.95 | -0.54 | 4210 | 2095 |

**TABLE 2:** MERS-CoV and SARS-CoV2 Gene Signatures Defined in this Study.

### 2.3 Identification of icSARS Associated Pathways

To identify pathway activity changes associated with SARS infection, we generated two icSARS pathway panels (Figure 2). To do this, we selected 500 pathways from the positive and negative tails from the GSE47960-derived icSARSvsmock pathway signature and used them to form two individual query pathway sets. We chose 500 pathways to capture maximum coverage of the signature that was allowable by GSEA [31]. GSEA compared each query pathway set to the GSE47961-derived icSARSvsmock pathway signature (reference). Leading-edge (LE) pathways from each analysis were collected, one group of pathways per tail. To make our pathway panels more robust, this process was reversed using GSE47960-derived icSARSvsmock pathway signature as reference and GSE47961-derived icSARSvsmock pathway signature for query pathway sets. We defined icSARS pathway panels from the pathways in common between leading-edge pathways from the same tail collected from both groups.

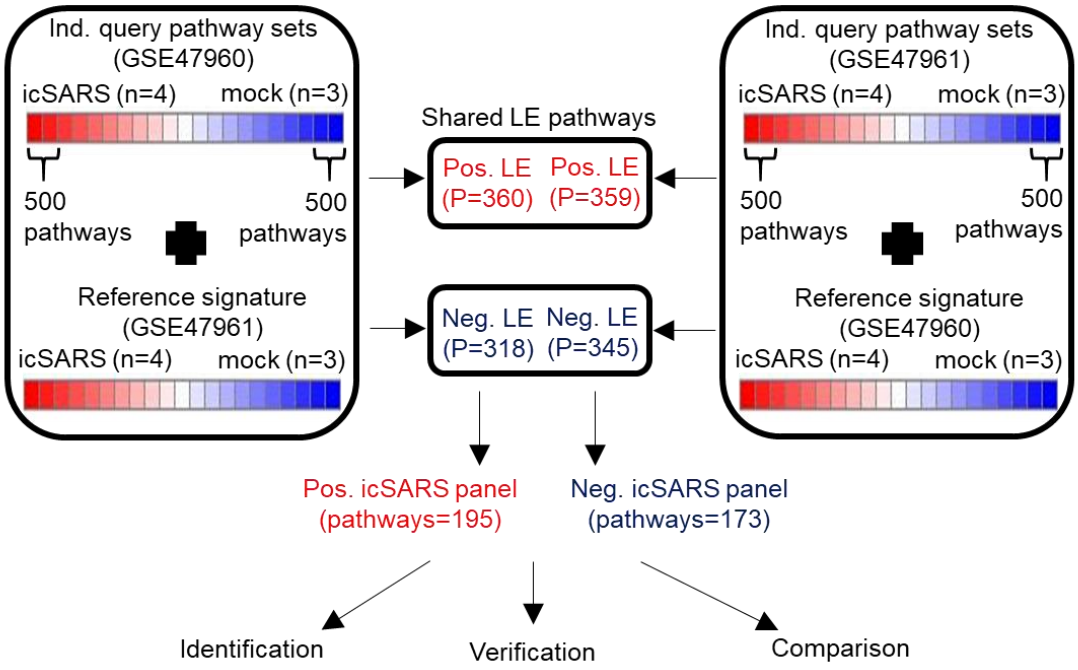


**FIGURE 2:** Generation of icSARS Pathway Panels.

### 2.4 Random Model Verification of icSARS Pathway Panel Enrichment

To verify pathway activity changes that we identified as associated with icSARS infection by inclusion in our icSARS pathway panels, we performed GSEA using icSARS pathway panels (individual queries) and six icSARSvsmock pathway signatures derived from human lung cultures, including the GSE47960-derived and GSE47961-derived icSARSvsmock signatures from which we generated the icSARS panels (individual references). To assess if results generated from GSEA could be achieved randomly, we randomly selected 1000 pathway panels consisting of 184-pathways from the 7573 GO BPs used to define pathway signatures (queries) for GSEA against 37 pathway signatures defined earlier (references). We selected 184 pathways for the query set size since this was the average icSARS pathway panel size, and our previous gene expression study observed no substantial difference in range of achieved NES generated from this random model approach when the query set size changed [35]. These analyses generated a null distribution of NES to which we compared the NES achieved by icSARS pathway panels for each reference pathway signature and count the number of equal or better NES to estimate significance (*i.e.*, distribution p-value). Box and whiskers plots were calculated using XLStat version 2020.3 [41].

### 2.5 Comparison of icSARS-induced Pathway Activity Changes to Changes in Specific SARS Strains and Models

We then extended our analysis to compare icSARS pathway panels to 19 pathway signatures derived from human lung cultures infected with other SARS strains (Urbani, BatSRBD, ΔORF6, ΔNSP16, ExoNI, icMERS or SARS-CoV2). To examine if the pathway activity changes associated with SARS infection in human lung cultures also were associated with infection in infected mice, we used GSEA to compare icSARS pathway panels to 12 pathway signatures derived from lung samples of mice infected with icSARS, MA15, ΔORF6, BatSRBD, or ΔNSP16. Heat maps were generated by Morpheus, https://software.broadinstitute.org/morpheus.

### 2.6 Meta-analysis Across SARS Strains and Models

To identify pathways generally associated with SARS infection, we collected same-tailed, leading-edge pathways generated by GSEA using our icSARS panel, and performed a meta-analysis retaining pathways in common across leading-edges. Pathways associated with SARS infection were identified by pathway membership across leading-edges identified in these 37 analyses. Pathways were not excluded from further consideration in our meta-analysis if they were not included in a leading-edge due to lack of representation in that dataset's platform. To determine which identified pathways were most associated with SARS infection, Stouffer's z scores were calculated using GSEA p-values from all pathway signature comparisons of selected pathways using the stats.combine_pvalues command from the scipy package in Python 3.8.

## 3. RESULTS

### 3.1 Meta-analysis Across SARS Strains and Models

To identify pathways associated with response to an icSARS infection, we began by defining the GSE47960-derived icSARSvsmock and GSE47961-derived icSARSvsmock pathway signatures (Table 1) We used these signatures to generate four pathway sets, each containing the 500 most differentially active pathways from the positive or negative tails of GSE47960-derived icSARSvsmock (NES>1.3 and <-1.1 for positive and negative tails, respectively) or GSE47961-derived icSARSvsmock (NES>1.3 and <-1.3 for positive and negative tails, respectively). To assess similarity between these two pathway signatures, we first calculated enrichment using GSEA between either GSE47960-derived icSARSvsmock positive or negative tail pathway sets and the GSE47961-derived icSARSvsmock and achieved NES=5.50 and NES= 5.17 for positive and negative tail query pathway sets, respectively, both with a GSEA p-value<0.001. We observed similar findings when GSEA calculated enrichment between either GSE47961-derived icSARSvsmock positive (NES=6.0) or negative (NES= 5.0) tail pathway sets and the GSE47960-derived icSARSvsmock (p-value<0.001). We defined separate positive and negative icSARS pathway panels from common leading-edge pathways identified from analyses (Supplemental Material STable 2), representing up- and down-regulated pathways associated with icSARS infection.

Among pathways in positive icSARS panel, there were 100 pathways (51.0% of the panel) associated with regulation of cellular processes and 26 pathways (13.3% of the panel) associated with cellular response. We noted 71 pathways (36.2% of the panel) associated with the immune system or viral processes with 29 of those pathways (14.8% of the panel) not directly related to regulation or cellular response (*e.g.*, neutrophil homeostasis, GO BP 0001780). Of the positive icSARS panel pathways, 17 pathways were identified by our prior gene expression meta-analysis [35]. Those pathways included negative regulation of viral genome replication (GO BP 0045071), response to type I IFN (GO BP 0034340), defense response to virus (GO BP 0051607), response to virus (GO BP 0009615), circadian regulation of gene expression (GO BP 0032922), circadian rhythm (GO BP 0007623), negative regulation of viral transcription (GO BP 0032897), positive regulation of PRI miRNA transcription by RNA polymerase II (GO BP 1902895), response to muscle stretch (GO BP 0035994), negative regulation of DNA binding (GO BP 0043392), epithelial tube branching involved in lung morphogenesis (GO BP 0060441), negative regulation of DNA binding transcription factor activity (GO BP 0043433), response to cAMP (GO BP 0051591), positive regulation of p38MAPK cascade (GO BP 1900745), response to

corticosterone (GO BP 0051412), positive regulation of DNA binding transcription factor activity (GO BP 0051091), and positive regulation of NF-κB transcription factor activity (GO BP 0051092). These results demonstrated the predictive ability of our pathway signature approach. In the negative icSARS pathway panel, we found 54 pathways (31.0% of the panel) associated with regulation of cellular processes, 19 pathways (10.1% of the panel) related to ciliary action, 28 pathways (16.1% of the panel) associated with nucleotide and protein production and modification, and 29 pathways (16.7% of the panel) connected to either mitotic or meiotic cell division. Of the negative icSARS panel pathways, no pathways were identified by our prior gene expression meta-analysis [35]. We speculated that these identified pathways without previously reported associations with icSARS infections also were associated with an icSARS infection.

To assess if our pathway signature approach was better able to identify pathway activity changes associated with icSARS infection than our previous gene signature approach [35], we compared icSARS panel pathways to pathways identified by FET previously. Our current pathway signature analysis identified 386 pathways with a nominal p-value<0.05 (Supplemental Material STable 2) whereas FET found only 59 significant GO-BP pathways (EASE score p-value<0.05) from icSARS gene panel [35]. We found 17 pathways in common between these analyses: negative regulation of viral genome replication (GO BP 0045071), response to type I IFN (GO BP 0034340), defense response to virus (GO BP 0051607), response to virus (GO BP 0009615), circadian regulation of gene expression (GO BP 0032922), circadian rhythm (GO BP 0007623), negative regulation of viral transcription (GO BP 0032897), response to muscle stretch (GO BP 0035994), negative regulation of DNA binding (GO BP 0043392), epithelial tube branching involved in lung morphogenesis (GO BP 0060441), negative regulation of DNA binding transcription factor activity (GO BP 0043433), response to cAMP (GO BP 0051591), positive regulation of p38MAPK cascade (GO BP 1900745), response to corticosterone (GO BP 0051412), positive regulation of DNA binding transcription factor activity (GO BP 0051091), and positive regulation of NF-κB transcription factor activity (GO BP 0051092) and positive regulation of PRI miRNA transcription by RNA polymerase II (GO BP 1902895). These results demonstrate the expanded ability of pathway signatures to identify established and new pathway activity changes associated with SARS infection.

### 3.2 Enrichment of icSARS Pathway Panels and Specific icSARS Panel Pathways Verified in Independent Datasets

To verify our icSARS pathway panels were associated with response to an icSARS infection, we used GSEA to calculate enrichment between our icSARS panels and four icSARSvsmock verification pathway signatures from human cell cultures (Table 1). We began by performing GSEA between our icSARS panels and the two identification icSARSvsmock signatures used to define the panels to determine the potential range of achievable NES for the panels. We found NES from the identification signatures were >8 or <-8 for positive and negative icSARS panels, respectively, all with GSEA p-value<0.001 (Figure 3A). GSE47962 was conducted using the same experimental conditions as the two identification datasets, so GSE47962 is a true validation dataset for our icSARS pathway panels. Therefore, we compared the icSARS pathway panels to the GSE47962-derived pathway signature and achieved significant enrichment (NES=5.44 and -5.56, both GSEA p-value<0.001) for both panels (Figure 3B). We then repeated GSEA with our icSARS pathway panels and three other verification signatures derived from experiments using a different lung culture cell type. We observed significant enrichment (NES > 5 for the positive icSARS panel and <-4 for the negative icSARS panel, Figure 3C, all with GSEA p-value<0.001) like what we observed with GSE47962. To determine how likely the NES achieved for icSARS panels would be achieved by random chance, we generated 1000 randomly selected 184-pathway panels from shared pathways between the GSE47960-derived and GSE47961-derived icSARSvsmock signatures to match the potential composition of our icSARS panels. We then repeated GSEA using these randomly generated pathway panels and identification and verification signatures to generate a null distribution of NES achieved via random chance. From this, we found random NES ranged from 2.08 to -2.20 (Figure 3D) and there was no distinction in NES range among icSARSvsmock signatures, illustrating that NES achieved by our icSARS

panels are non-random (null distribution p-value<0.001). Taken together, we concluded that the enrichment achieved from our icSARS panels was true and reproducible.

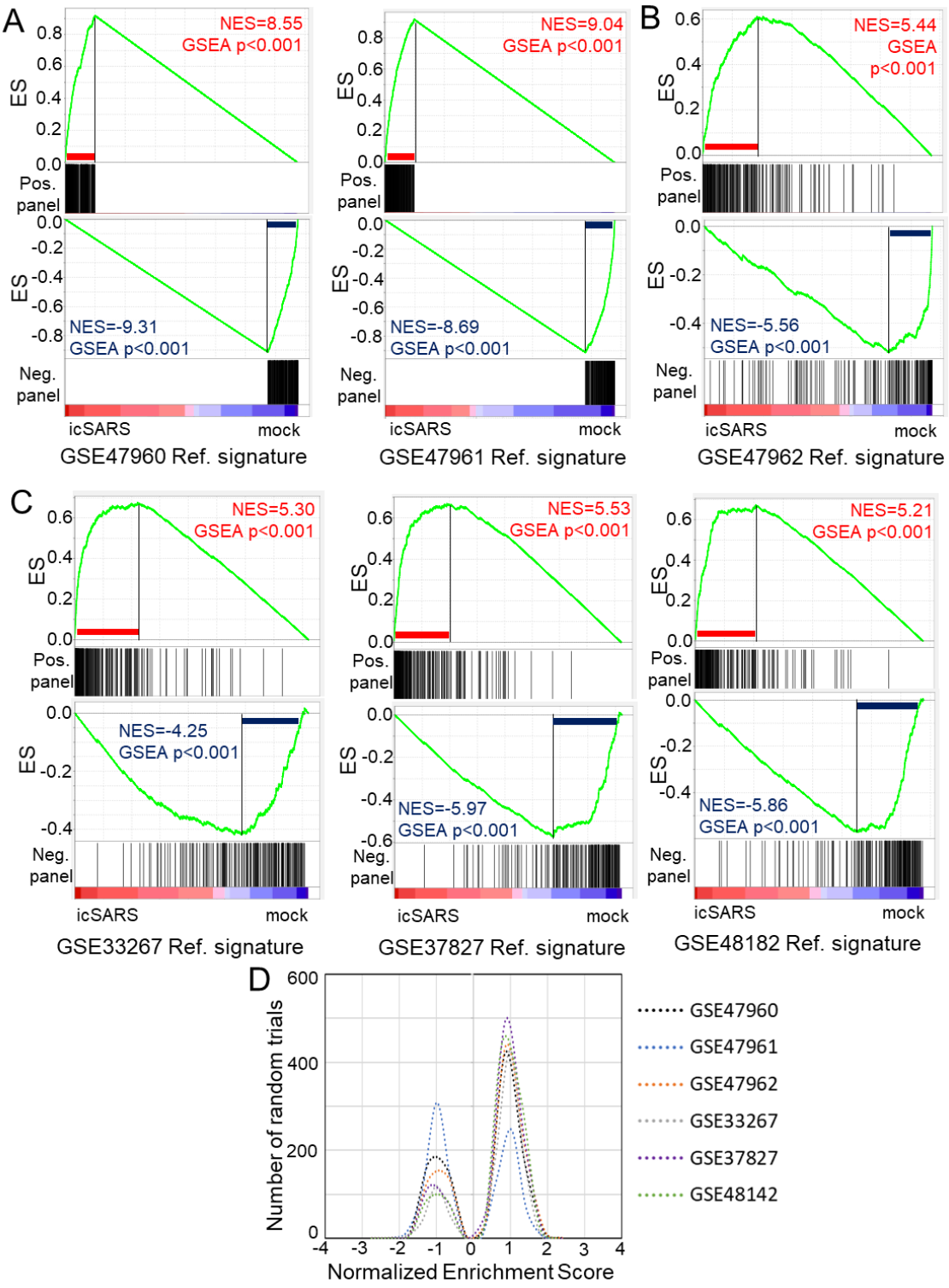


**FIGURE 3:** Verification of icSARS Pathway Panels in Independent Datasets.

To determine which icSARS panel pathways were verified across all signatures, we examined leading-edge pathways identified by GSEA for each identification and verification signature. Across six icSARSvsmock signatures in human cell cultures, 115 pathways were included in all leading-edges (Supplemental Materials STable 3). Of the positive icSARS panel pathways, 13 pathways were identified by our prior gene expression meta-analysis [35]. Those pathways included negative regulation of viral genome replication (GO BP 0045071), response to type I IFN (GO BP 0034340), defense response to virus (GO BP 0051607), response to virus (GO BP 0009615), circadian regulation of gene expression (GO BP 0032922), circadian rhythm (GO BP 0007623), negative regulation of viral transcription (GO BP 0032897), positive regulation of PRI miRNA transcription by RNA polymerase II (GO BP 1902895), response to cAMP (GO BP 0051591), positive regulation of p38MAPK cascade (GO BP 1900745), response to corticosterone (GO BP 0051412), positive regulation of DNA binding transcription factor activity (GO BP 0051091), and positive regulation of NF-κB transcription factor activity (GO BP 0051092). These data supported the conclusion that our shared leading-edge pathways were associated with icSARS infection in human lung cultures and supported the hypothesis that identified pathways without previously reported associations also were involved in icSARS infection in human lung cultures.

### 3.3 Positive icSARS Pathway Panel Significantly Enriched Across Models and SARS Strains

To determine if icSARS pathway panels were associated with response to an icSARS infection in vivo, we extended our analysis to a mouse model by using GSEA on icSARS pathway panels and an icSARSvsmock signature derived from mouse lung samples (Table 1). We found significant enrichment for the positive icSARS panel (NES=3.99, GSEA p-value<0.001, Figure 4A) that was non-random (null distribution p-value<0.001, Figure 4B). However, for the negative icSARS panel, we observed positive enrichment of the negative icSARS panel (NES=1.45, p-value=0.041), indicating that the negative icSARS panel was not reproducible in a mouse model. Taken together, these results demonstrated that the enrichment achieved from our positive icSARS panel was true for icSARS infection, but the same is not true for our negative icSARS panel.

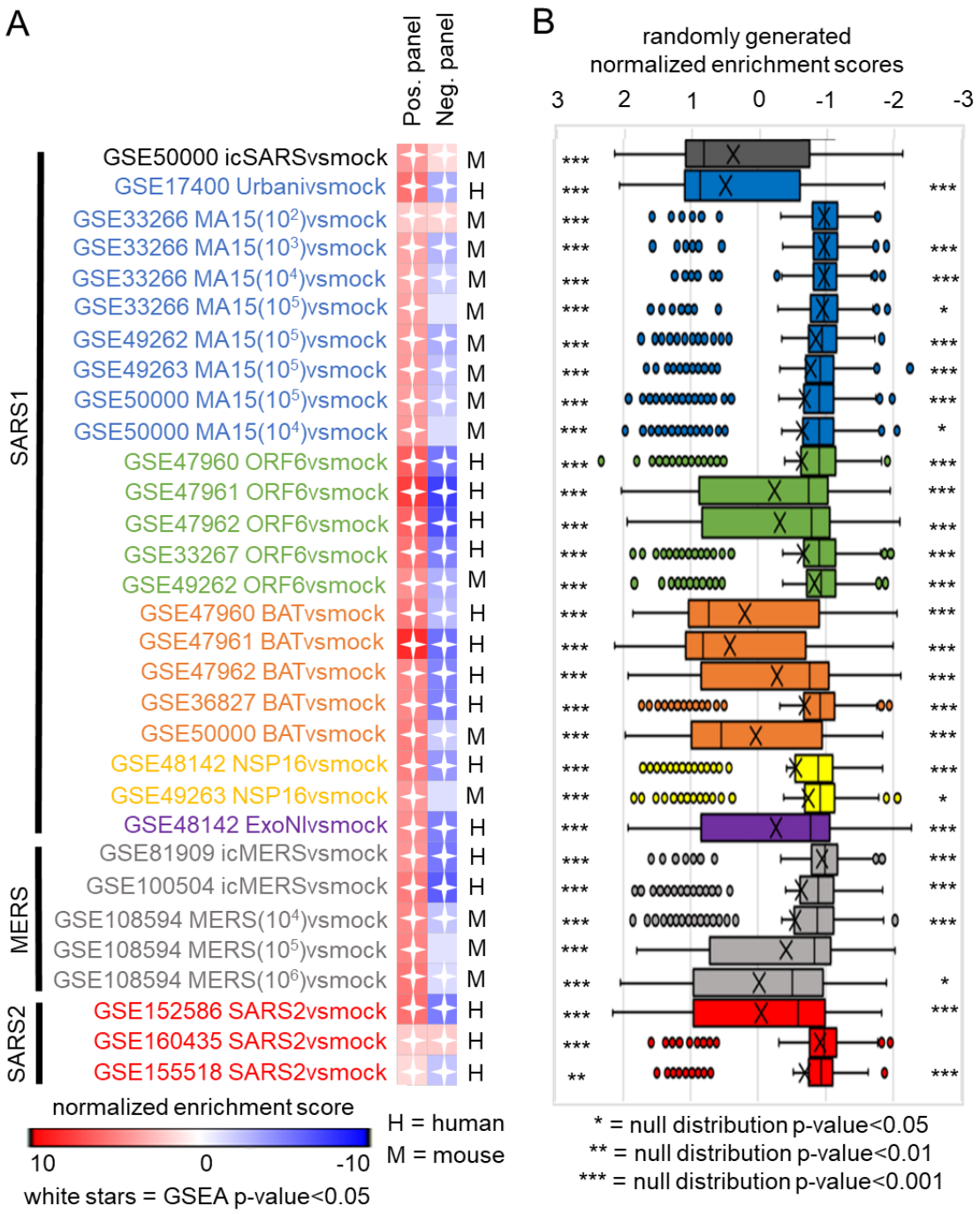


**FIGURE 4:** Non-random Enrichment of Positive icSARS Pathway Panel Across SARS Infected Cultures and Samples Revealed Pathway Activity Changes Associated with SARS Infection.

Next, we expanded our analysis to include other SARS-CoV1 strains, starting with the other fully virulent strains, Urbani and its mouse adapted strain MA15 (Table 1). For human cell cultures, we found non-random (NES range 1.86 to -2.05, null distribution p-value<0.001, Figure 4B) enrichment between the Urbanivsmock pathway signature and both positive (NES=5.52, GSEA p-value<0.001) and negative (NES= 3.16, p-value<0.001) icSARS pathway panels (Figure 4A). For mouse samples, we found consistent non-random (NES range 2.05 to -1.97, null distribution p-value<0.001, Figure 4B) enrichment for the positive icSARS panel (NES range 4.03 to 2.16,

null distribution p-value<0.001, Figure 4A). We found inconsistent enrichment for the negative icSARS panel, with the GSE33266-derived MA15($10^2$)vsmock pathway signature achieving significant positive enrichment (NES=1.72, p-value=0.001), similar to what we previously observed in the icSARSvsmock signature from mice.

We then examined pathway activity in attenuated SARS-CoV1 strains (ΔORF6, BatSRBD, ΔNSP16, and ExoNI, Table 1). We again found consistent enrichment despite model (*e.g.*, human or mouse) when comparing the positive icSARS panel to pathway signatures derived from attenuated SARS-CoV1 strains, achieving NES ranging from 8.32 to 4.00 with GSEA p-value<0.001 (Figure 4A) that was non-random (NES range 2.27 to -2.32, null distribution p-value<0.001, Figure 4B). As seen with icSARS in the mouse model, inconsistent negative enrichment was observed for the negative icSARS panel. Taken together, these results demonstrated that the enrichment achieved from our positive icSARS panel, but not our negative icSARS panel, was consistent across infections of different SARS-CoV1 strains in both human and mouse models.

We further expanded our analysis to include other SARS strains, specifically MERS and SARS-CoV2 (Table 2). We again found consistent enrichment for the positive icSARS panel when performing GSEA on MERSvsmock and SARS2vsmock pathway signatures (NES ranging 5.60 to 1.54, GSEA p-value<0.005, Figure 4A) that was non-random (NES ranging 2.02 to -2.15, null distribution p-value<0.005, Figure 4B). However, inconsistent enrichment was observed again for the negative icSARS panel with GSE155518-derived SARS2vsmock achieving significant positive enrichment (NES=2.08, GSEA p-value<0.001) and GSE108594-derived MERS($10^5$)vsmock failing to achieve significant enrichment (NES= 1.14, p-value=0.249). Overall, these results demonstrated that the enrichment achieved from our icSARS positive panel was consistent across SARS infections regardless of strain in both human and mouse models.

### 3.4 Meta-analysis of Positive icSARS Panel Leading-edge Pathways Revealed Eleven Top Pathways

To determine which pathways were most relevant to SARS infection, we performed a meta-analysis of leading-edge pathways identified by GSEA with each SARS pathway signature and the positive icSARS pathway panel. Pathways identified through leading-edge intersections represent pathways associated with infection of that specific SARS strain. We then find shared leading-edge pathways across SARS strains to identify top pathway candidates to potentially target therapeutically. We chose not to perform the same analysis with the negative icSARS pathway panel due to its inability to consistently achieve non-random enrichment across SARS pathway signatures.

We began by performing meta-analysis across signatures between icSARS infected human and mouse models. For the icSARSvsmock signature from mice, 95 pathways were identified in the leading-edge and 65 of those pathways were shared with leading-edge pathways across all six signatures from human cell cultures (Supplemental Materials STable 4). We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to determine 66 pathways are associated with icSARS infection regardless of model (Figure 5A). For the eight MA15vsmock signatures from mice, 85 leading-edge pathways were shared across all eight signatures (Supplemental Materials STable 5). The Urbanivsmock signature from human cell cultures had 112 leading-edge pathways (Supplemental Materials STable 6) with 79 pathways in common with shared leading-edge pathways from MA15vsmock signatures. We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to detect a total of 88 pathways (Figure 5B). Across five ORF6vsmock signatures in human cell cultures, 111 pathways were included in all leading-edges (Supplemental Materials STable 7). For the icSARSvsmock signature from mice, 129 pathways were identified in the leading-edge with 83 pathways shared with human cell cultures (Supplemental Materials STable 8). We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to identify a total of 84 pathways (Figure 5C). Across five BATvsmock signatures in

human cell cultures, 71 pathways were included in all leading-edges (Supplemental Materials STable 9). For the one BATvsmock signature from mice, 129 pathways were identified in the leading-edge with 83 pathways shared with human cell cultures (Supplemental Materials STable 10). We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to find a total of 59 pathways (Figure 5D). There were two NSP16vsmock signatures, one in human cell cultures and the other in mice. From these, 153 pathways were included in the human-derived leading-edge (Supplemental Materials STable 11) and 120 pathways were identified in the mouse-derived leading-edge (Supplemental Materials STable 12) with 115 shared between the two leading-edges. We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to reveal a total of 117 pathways (Figure 5E). For the ExoNIvsmock signature, 106 pathways were identified in the leading-edge (Supplemental Materials STable 13). Across two MERSvsmock signatures in human cell cultures, 126 pathways were included in all leading-edges (Supplemental Materials STable 14). For the three MERSvsmock signatures from mice, 105 pathways were included in all leading-edges with 91 pathways shared with human cell cultures (Supplemental Materials STable 15). We further considered leading-edge pathways that were consistently identified by GSEA when sufficiently represented in the dataset's platform to find a total of 93 pathways (Figure 5F). For SARS-CoV2, we found 28 pathways shared across three SARS2vsmock signatures (Supplemental Materials STable 16). Supplemental Materials STable 17 details pathway inclusion for each dataset specifically.

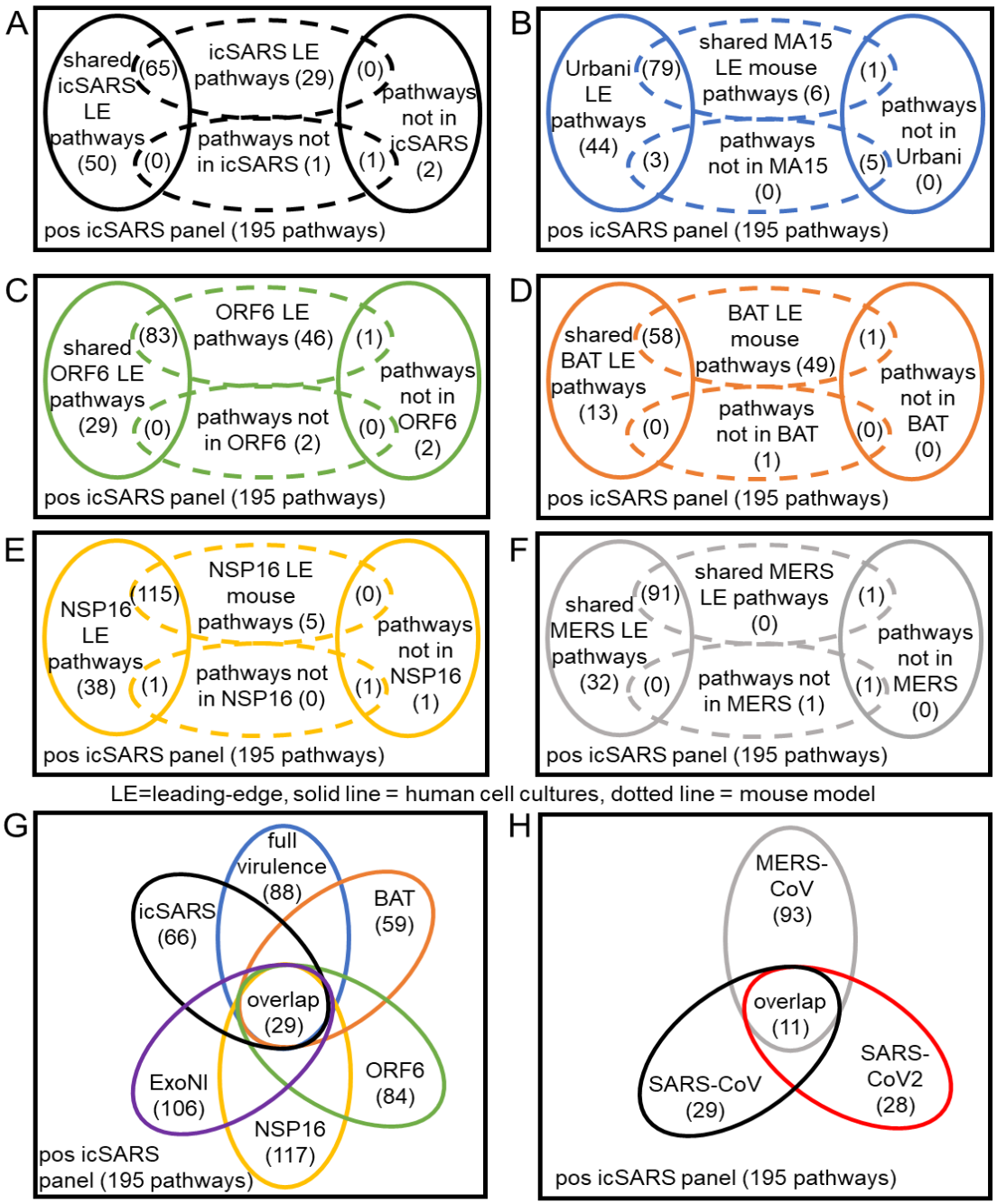


**FIGURE 5:** Meta-analysis of Leading-edge Positive icSARS Panel Pathways Across Pathway Signatures Revealed Eleven Up-regulated Pathways Associated with SARS Infection.

Now that pathways associated with individual SARS strains had been identified across models, we performed meta-analysis on these pathway lists across SARS strains to find pathways associated with SARS infection generally. Across 29 SARS-CoV1 pathway signatures spanning seven strains with varying levels of virulence and inoculating doses, we found 29 pathways associated with SARS-CoV1 infection (Figure 5G). When considering leading-edge pathways from the five MERS-CoV and three SARS-CoV2 pathway signatures, we found 11 pathways were shared across MERS-CoV and SARS-CoV2 infections also (Figure 5H). The 11 pathways were described briefly in Table 3 with more detail found in Supplemental Materials STable 1.

| GO BP | Pathway Name | Pathway Description | Number of Genes |
|---|---|---|---|
| 2000516 | Positive regulation of CD4 positive alpha beta T-cell activation | Any process that activates or increases the frequency, rate, or extent of CD4-positive, alpha-beta T cell activation. | 36 |
| 0034340 | Response to type I interferon | Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) because of a type I interferon stimulus. | 99 |
| 0002286 | T-cell activation involved in immune response | The change in morphology and behavior of a mature or immature T cell resulting from exposure to a mitogen, cytokine, chemokine, cellular ligand, or an antigen for which it is specific, leading to the initiation or perpetuation of an immune response. | 109 |
| 0043367 | CD4 positive alpha beta T-cell differentiation | The process in which a relatively unspecialized T cell acquires specialized features of a mature CD4-positive, alpha-beta T cell. | 78 |
| 0001818 | Negative regulation of cytokine production | Any process that stops, prevents, or reduces the rate of production of a cytokine. | 368 |
| 0002292 | T-cell differentiation involved in immune response | The process in which an antigenically naive T cell acquires the specialized features of an effector, regulatory, or memory T cell as part of an immune response. Effector T cells include cells which provide T cell help or exhibit cytotoxicity towards other cells. | 71 |
| 0002287 | Alpha beta T-cell activation involved in immune response | The change in morphology and behavior of an alpha-beta T cell resulting from exposure to a mitogen, cytokine, chemokine, cellular ligand, or an antigen for which it is specific, leading to the initiation or perpetuation of an immune response. | 63 |
| 2000514 | Regulation of CD4 positive alpha beta T-cell activation | Any process that modulates the frequency, rate or extent of CD4-positive, alpha-beta T cell activation. | 65 |
| 0042092 | Type 2 immune response | An immune response which is associated with resistance to extracellular organisms and pathological conditions, which is orchestrated by the production of cytokines by any of a variety of cell types resulting in enhanced production of certain antibody isotypes and other effects. | 39 |
| 0046631 | Alpha beta T-cell activation | The change in morphology and behavior of an alpha-beta T cell resulting from exposure to a mitogen, cytokine, chemokine, cellular ligand, or an antigen for which it is specific. | 146 |
| 0035710 | CD4 positive alpha beta T-cell activation | The change in morphology and behavior of a CD4-positive, alpha-beta T cell resulting from exposure to a mitogen, cytokine, chemokine, cellular ligand, or an antigen for which it is specific. | 97 |

**TABLE 3:** Top Pathways Associated with SARS Infection Identified in this Study.

### 3.5 Type I Interferon Pathway Identified as Most Associated with SARS Infection

To determine which of the 11 identified pathways were the most associated with SARS infection, we collected p-values generated by GSEA for each pathway and signature. From these p-values, we performed Stouffer's z score analysis for combined p-values. We found that the response to type I IFN pathway had the highest Stouffer's z-score (Figure 6), indicating that pathway was the most associated with SARS infection. We also noticed that all pathways had at least one signature where GSEA p-value>0.05, highlighting the advantage of a meta-analysis approach reliant on enrichment (*i.e.*, leading-edge membership) rather than statistical significance (*i.e.*, p-value). When comparing these 11 pathways to those previously identified by our prior gene expression meta-analysis [35], we found only one icSARS panel pathway, response to type I IFN, in common. These findings demonstrate the ability of our pathway signature meta-analysis approach in identifying pathway activity changes not previously identified via Fisher's Exact Test methods.

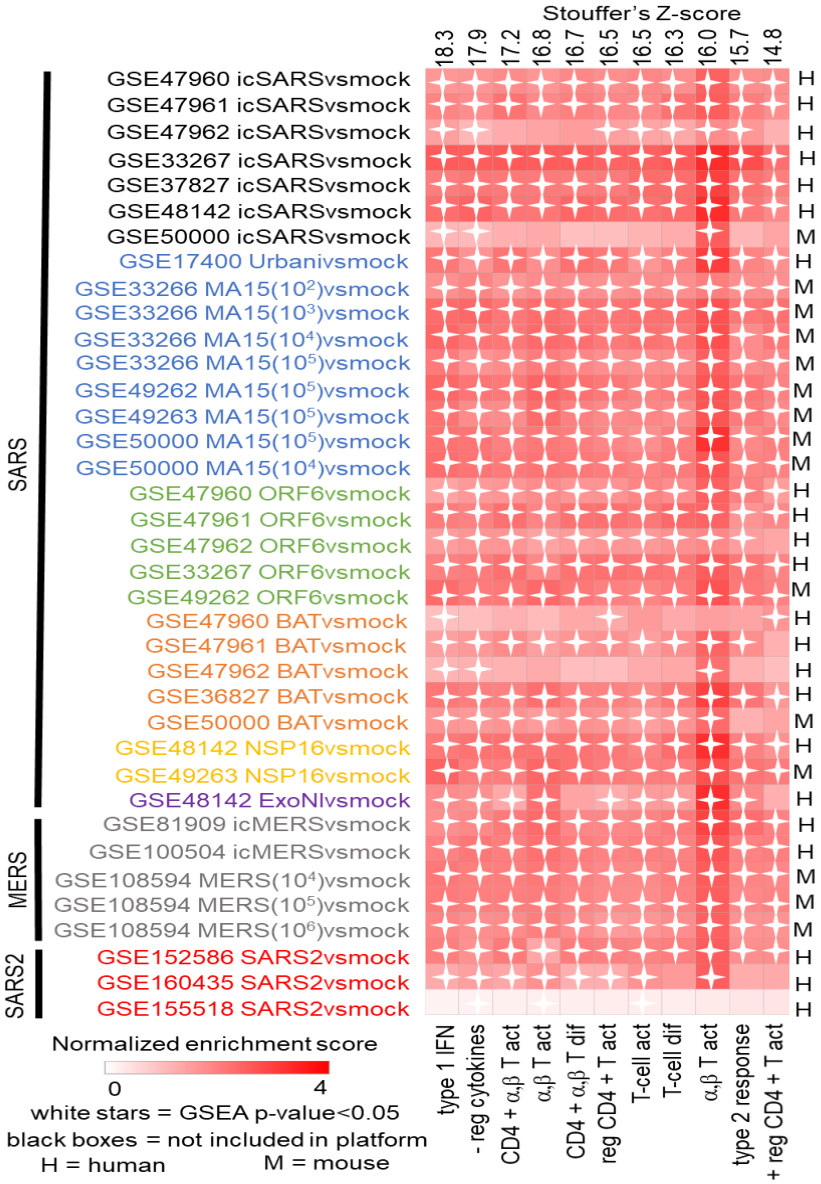


**FIGURE 6:** Enrichment examination for Eleven Identified Pathways Reveals Type I IFN as Most Consistently Up-regulated Across SARS Strains.

## 4. DISCUSSION

SARS infections remain a serious public health threat due to their strong pandemic causing potential. While efforts have gone into developing effective therapeutics for SARS-infected patients, treatment options are still limited, due in part to an incomplete understanding of the molecular changes driving SARS infections. Identification of differentially regulated pathways associated with SARS infections can improve our understanding of SARS-induced molecular changes, potentially contributing to the development of new therapeutic options to use in the fight against SARS and future SARS outbreaks. This work performed a GSEA-based, meta-analysis of pathway signatures generated from mRNA expression data across SARS-CoV1, MERS-CoV, and SARS-CoV2 infections to reveal differentially regulated pathways associated with SARS infections.

Our approach identified response to type I IFN as the top of 11 immune system pathways found consistently across SARS strains (Figure 6). Type I IFNs are a large group of proteins that assist in regulating innate immune system activity and activating adaptive immunity [42, 43]. Type I IFNs are directly connected with immune response toward a wide variety of bacterial and viral infections, including hepatitis C virus, human immunodeficiency virus, varicella-zoster virus, and SARS-CoV2 [42, 44]. Prior reports demonstrated the release of pro-inflammatory cytokines, especially IFN-α and IFN-γ, was correlated with lethality in SARS patients [45]. Other reports found that levels of IFN-γ were lower in CD4+ T-cells from patients with severe versus mild symptoms and suggested that the infection may initially affect CD4+ and CD8+ T-cells, reducing the production of IFN-γ [46]. This work here is not the first time that computational analysis predicted the involvement of type I IFNs in SARS infection. Our previous examination of gene-expression changes associated with icSARS infections identified the response to type I IFNs pathway through enrichment analysis using a Fisher's Exact Test variant [35]. GSEA-based meta-analysis of the gene signatures we used to define pathway signatures used in this study found five genes involved in type I IFN response. We were not surprised to find four of those five identified genes (XAF1, OASL, OAS3, and IFIT3) included among the 99 genes used by the GO BP knowledgebase to define the response to type I IFNs pathway (Supplemental Material STable 1). Therefore, we concluded that targeting IFN response therapeutically might improve outcomes for patients with SARS infections. Our conclusion supports several reports that examined the success of IFN therapy for SARS-CoV2 patients with promising results though details of therapeutic timing, type of IFN, and administration route are still highly debated [12-14, 18].

While the connection between Type I IFNs and SARS-CoV2 infections is well established, our GSEA-based meta-analysis approach improves upon our prior gene signature method by also identifying 10 immune system pathways that were not previously identified (Table 3). Those pathways included the type 2 immune response pathway, the negative regulation of cytokine production, and eight pathways involved in T-cell activation or differentiation with four pathways specifically related to CD4+ immune cells. We were not surprised by these findings due to the extensive literature documenting the cytokine storm classically associated with SARS-CoV2 infections [47-51]. The type 2 immune response is an adaptive response where differentiated T helper (CD4+) cells drive eosinophil recruitment, maturation of killer T cells (CD8+), and antibody production via secretion of a distinct selection of cytokines including IL-4, IL-5, and IL-13 [52, 53]. Previous reports have shown CD4+ T cells are activated from SARS-CoV2 infection as seen by increased expression of cellular markers like HLA-DR, CD25, CD38 and Ki-67 [54]. While these results demonstrate the predictability of our approach, many pathways with reported connections to SARS were not detected as top pathway candidates. For example, no pathways related to CD8+ cells were found despite one pathway (CD8_POSITIVE_ALPHA_BETA_T_CELL_ACTIVATION, GO:0036037) being included in the positive icSARS pathway panel (Supplemental Material STable 2). Another example is NF-κB signaling which has reported associations with SARS-CoV2 infection and pharmacological inactivation of the phosphorylation of inhibitor of NF-κB kinase subunit beta has shown promise as an effective treatment for COVID-19 symptoms [55-57]. There were eight NF-κB pathways in

our defined positive icSARS pathway panel (Supplemental Material STable 2), but none were detected among top pathway candidates (Table 3). While our meta-analysis approach was able to detect known pathways associated with SARS infection, clearly further efforts are needed to optimize our approach's detection ability.

In this study, we observed some interesting pathway detection threshold findings that may have biological implications. We noted that consistent enrichment of the negative icSARS pathway panel could not be achieved across SARS strains and models (Figure 4). This finding was not surprising since we also could not achieve consistent enrichment of the negative icSARS gene panel across SARS strains and models in our prior GSEA-based gene expression meta-analysis [35]. This observation may have biological implications since several reports show characteristic changes in gene expression when cells of different types are infected with SARS-CoV2 and could be related to the intense immune response seen in some COVID-19 patients [23, 30, 58-60]. We also noted that our pathway meta-analysis approach was able to calculate enrichment between the positive icSARS pathway panel and the positive tail of the pathway signature with the lowest inoculation dose, GSE33266 MA15($10^2$)vsmock, which is better that what we observed for the GSE33266 MA15($10^2$)vsmock gene signature in our gene expression meta-analysis study [35]. This finding suggests that use of pathway signatures may be better than gene signatures at detecting molecular changes early in a patient's infection, an improvement over our prior approach.

While this meta-analysis revealed pathways with already well-established associations to SARS infections, the work has some limitations. Our study is limited by the need of our GSEA-based meta-analysis approach for control samples to establish differential activity for pathway signature definition. The requirement for mock-derived data limits the ability to use our approach on other datasets and limited us to the use of specific publicly available mRNA expression data. Datasets selected for this study contained human lung cultures or mouse lung samples due to their public availability at the time this study was conducted and focus on 48hr time point data only. This study could be enhanced if mRNA expression data from SARS and mock infected lung samples from human patients was available for examination. Currently mRNA expression data is available from autopsy lung samples from SARS patients only. Further, this purely bioinformatic work was limited by a lack of direct experimental evidence. We were unable to conduct follow-up experiments using other techniques, such as pathway reporter assays, Western blotting, ELISA, or qRT-PCR, to confirm our top pathway candidate predictions that were generated only from mRNA expression data. Experimental work will be a necessary step to implementing results obtained in this work.

Based on the results from this study, future directions should examine the potential of IFN as a therapeutic option for SARS infections. For example, gene expression data from IFN-treated and untreated SARS-CoV2 human cell cultures or mouse lung samples would be of particular interest for future studies. We predict IFNtreatedSARSvsuntreatedSARS pathway signatures would be reversed (*i.e.*, positive icSARS panel achieves significant enrichment with a negative NES), indicating IFN-treated samples looked more like untreated cultures and samples from this study, supporting the conclusion of targeting IFN as viable therapeutic option for SARS infections. Future directions also could include a more in-depth computational analysis of gene inclusion and overlap among pathways identified here. Such an analysis could improve our understanding of mechanisms behind SARS infections which could lead to a more specific determination of molecular targets for therapeutic development.

This work applied a GSEA-based, meta-analysis approach for analyzing pathway signatures from gene expression data to determine if such an approach would overcome FET limitations and identify more pathways associated with SARS infections than observed in our previous work using gene signatures. By using pathway signatures, we confirmed our prior gene expression findings by identifying response to type I IFN as the top pathway candidate most associated with SARS infection and expanded upon our prior gene expression findings by identifying 10 consistently up-regulated pathways also associated with SARS infection that were not previously

identified using a FET-variant. Our results support prior reports that targeting type I IFN response as a solo or co-therapy could be an effective therapy for SARS infections as IFN therapy has been noted to successfully treat patients with moderate COVID-19 infections. Further, our results identified immune system pathways involved in T-cell activation and differentiation with a specific focus on CD4+ immune cell activity that were not previously identified using FET. These pathways could be further examined to identify new targets for treatment of COVID-19 patients. Overall, this study displayed our pathway signature approach's ability to consistently detect new and established biologically important pathways and its potential as a beneficial computational approach to discover new molecular insights with therapeutic promise that can directly impact how clinicians treat COVID-19 infections and improve COVID-19 patient outcomes.

## 5. ACKNOWLEDGEMENTS

The authors would like to thank Terri Pulice for the graphic design assistance, Johnathan Kennedy for the gene analysis on identified pathways, and Anton Oestereicher for the ongoing financial support.

## 6. REFERENCES

[1] X. Li, H.K.H. Luk, S.K.P. Lau, and P.C.Y Woo. (2019, Mar.) "Human Coronaviruses: General Features." Reference Module in Biomedical Sciences. [On-line]. Available: https://doi.org/10.1016/B978-0-12-801238-3.95704-0 [June 25, 2021].

[2] N. Petrosillo, G. Viceconte, O. Ergonul, G. Ippolito, and E. Petersen. "COVID-19, SARS and MERS: are they closely related?" Clinical Microbiology and Infection, vol. 26(6), pp. 729-734, Mar. 2020.

[3] P. Krishnamoorthy, A.S. Raj, S. Roy, N.S. Kumar, and H. Kumar. (2021, Jan.) "Comparative transcriptome analysis of SARS-CoV, MERS-CoV, and SARS-CoV-2 to identify potential pathways for drug repurposing." Computers in Biology and Medicine. [On-line]. 128. Available: 10.1016/j.compbiomed.2020.104123 [June 25, 2021].

[4] J.Y. Chung, M.N. Thone, and Y.J. Kwon. "COVID-19 vaccines: The status and perspectives in delivery points of view." Advanced Drug Delivery Reviews, vol.170, pp.1-25, Mar. 2021.

[5] E. Dong, H. Du, and L. Gardner. "An interactive web-based dashboard to track COVID-19 in real time". The Lancet Infectious Diseases, vol. 20(5), pp. 533-534, May 2020.

[6] S. Bahadur, W. Long, and M. Shuaib. "Human coronaviruses with emphasis on the COVID-19 outbreak." Virusdisease, vol. 31(2), pp. 80-84, Jun. 2020.

[7] M. Bhattacharya, A.R. Sharma, B. Mallick, G. Sharma, S.S. Lee, and C. Chakraborty. "Immunoinformatics approach to understand molecular interaction between multi-epitopic regions of SARS-CoV-2 spike-protein with TLR4/MD-2 complex." Infection, Genetics and Evolution. [On-line]. 85. Available: 10.1016/j.meegid.2020.104587.

[8] R.S. Baric. "Emergence of a Highly Fit SARS-CoV-2 Variant." New England Journal of Medicine, vol. 383(27), pp. 2684-2686, Dec. 2020.

[9] Y. Wu, W. Ho, Y. Huang, D.Y. Jin, S. Li, S.L. Liu, X. Liu, J. Qiu, Y. Sang, Q. Wang, K.Y. Yuen, and Z.M. Zhenglet. "SARS-CoV-2 is an appropriate name for the new coronavirus." Lancet, vol 395(10228), pp. 949-950, Mar. 2020.

[10] T. Kirby. "New variant of SARS-CoV-2 in UK causes surge of COVID-19." The Lancet Respiratory Medicine, vol. 9(2), pp. e20-e21, Feb. 2021.

[11] Y. Yi, P.N.P. Lagniton, S. Ye, E. Li, and R.H. Xu. “COVID-19: what has been learned and to be learned about the novel coronavirus disease.” International Journal of Biological Sciences, 16(10), pp. 1753-1766, Mar. 2020.

[12] M. Haji Abdolvahab, S. Moradi-Kalbolandi, D. Zarei M, Bose, A.K. Majidzadeh, and L. Farahmand. “Potential role of interferons in treating COVID-19 patients.” International Immunopharmacology. [On-line]. 90. Available: https://www.sciencedirect.com/science/article/pii/S1567576920336389?via%3Dihub [June 25, 2021].

[13] H. Shuai, H. Chu, Y. Hou, D. Yang, Y. Wang, B. Hu, X. Huang, X. Zhang, Y. Chai, J.P. Cai, J.F. Chan, and K. Yuen. “Differential immune activation profile of SARS-CoV-2 and SARS-CoV infection in human lung and intestinal cells: Implications for treatment with IFN-beta and IFN inducer.” Journal of Infection, vol. 81(4), pp. e1-e10, Oct. 2020.

[14] L.H. Calabrese, T. Lenfant, and C. Calabrese. “Interferon therapy for COVID-19 and emerging infections: Prospects and concerns.” Cleveland Clinic Journal of Medicine. [On-line]. Available: https://www.ccjm.org/content/early/2020/12/01/ccjm.87a.ccc066.long [June 25, 2021].

[15] Y. Jiang, D. Chen, D. Cai, Y. Yi, and S. Jiang. “Effectiveness of remdesivir for the treatment of hospitalized COVID-19 persons: A network meta-analysis.” Journal of Medical Virology, vol. 93(2), pp. 1171-1174, Feb. 2021.

[16] J. Santos, S. Brierley, M.J. Gandhi, M.A. Cohen, P.C. Moschella, and A.B.L Declan. “Repurposing Therapeutics for Potential Treatment of SARS-CoV-2: A Review. Viruses.” Viruses. [On-line]. 12(7). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7412090/pdf/viruses-12-00705.pdf [June 25, 2021].

[17] S.K. Mishra and T. Tripathi. “One year update on the COVID-19 pandemic: Where are we now?” Acta Tropica. [On-line]. 214. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7695590/pdf/main.pdf [June 25, 2021].

[18] R.P. Saha, A.R. Sharma, M.K. Singh, S. Samanta, S. Bhakta, S. Mandal, M. Bhattacharya, S.S. Lee, and C. Chakraborty. “Repurposing Drugs, Ongoing Vaccine, and New Therapeutic Development Initiatives Against COVID-19.” Frontiers in Pharmacology. [On-line]. 11. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7466451/pdf/fphar-11-01258.pdf [June 25, 2021].

[19] C. Chakraborty, A.R. Sharma, G. Sharma, M. Bhattacharya, and S.S. Lee. “SARS-CoV-2 causing pneumonia-associated respiratory disorder (COVID-19): diagnostic and proposed therapeutic options.” European Review for Medical and Pharmacological Sciences, vol. 24(7), pp. 4016-4026, Apr. 2020.

[20] C. Chakraborty, A.R. Sharma, M. Bhattacharya, G. Sharma, S.S. Lee, and G. Agoramoorthy. “COVID-19: Consider IL-6 receptor antagonist for the therapy of cytokine storm syndrome in SARS-CoV-2 infected patients.” Journal of Medical Virology, vol. 92(11), pp. 2260-2262, May 2020.

[21] V. Pooladanda, S. Thatikonda, and C. Godugu. “The current understanding and potential therapeutic options to combat COVID-19.” Life Sciences. [On-line]. 254. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7207108/pdf/main.pdf [June 25, 2021].

[22] G. Agapito, C. Pastrello, and I. Jurisica. “Comprehensive pathway enrichment analysis workflows: COVID-19 case study.” Briefings in Bioinformatics, vol. 22(2), pp. 676-689, Dec. 2020.

[23] P.K. Jha, A. Vijay, A. Halu, S. Uchida, and M. Aikawa. 'Gene Expression Profiling Reveals the Shared and Distinct Transcriptional Signatures in Human Lung Epithelial Cells Infected With SARS-CoV-2, MERS-CoV, or SARS-CoV: Potential Implications in Cardiovascular Complications of COVID-19." Frontiers in Cardiovascular Medicine. [On-line]. 7. Available: https://www.frontiersin.org/articles/10.3389/fcvm.2020.623012/full [June 25, 2021].

[24] B. Vastrad, C. Vastrad, and A. Tengli. "Bioinformatics analyses of significant genes, related pathways, and candidate diagnostic biomarkers and molecular targets in SARS-CoV-2/COVID-19." Gene Reports. [On-line]. 21. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7854084/pdf/main.pdf [June 25, 2021].

[25] D. Blanco-Melo, B.E. Nilsson-Payant, W.C. Liu, S. Uhl, D. Hoagland, R. Moller, T.X. Jordan, K. Oishi, M. Panis, D. Sachs, T.T. Wang, R.E. Schwartz, J.K. Lim, R.A. Albrecht, and B.R. tenOever. "Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19." Cell, vol. 181(5), pp. 1036-1045 e1039, May 2020.

[26] V.D. Menachery, L.E. Gralinski, H.D. Mitchell, K.H. Dinnon, S.R. Leist, B.L. Yount, E.T. McAnarney, R.L. Graham, K.M. Waters, and R.S. Baric. (2018, Aug.). "Combination Attenuation Offers Strategy for Live Attenuated Coronavirus Vaccines." Journal of Virology. [On-line]. 92(17). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6096805/pdf/e00710-18.pdf [June 25, 2021].

[27] A. Islam and M.A. Khan. (2020, Nov.). "Lung transcriptome of a COVID-19 patient and systems biology predictions suggest impaired surfactant production which may be druggable by surfactant therapy." Scientific Reports. [On-line]. 10(1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7656460/pdf/41598_2020_Article_76404.pdf [June 25, 2021].

[28] T.A. Taz, K. Ahmed, B.K. Paul, F.A. Al-Zahrani, S.M.H. Mahmud, and M.A. Moni. "Identification of biomarkers and pathways for the SARS-CoV-2 infections that make complexities in pulmonary arterial hypertension patients." Briefings in Bioinformatics, vol. 22(2), pp. 1451-1465, Mar. 2021.

[29] Y.H. Wu, I.J. Yeh, N.N. Phan, M.C. Yen, H.L. Liu, C.Y. Wang, and H.P. Hsu. "Severe acute respiratory syndrome coronavirus (SARS-CoV)-2 infection induces dysregulation of immunity: in silico gene expression analysis." International Journal of Medical Sciences, vol. 18(5), pp. 1143-1152, Jan. 2021.

[30] A.R. Daamen, P. Bachali, K.A. Owen, K.M. Kingsmore, E.L. Hubbard, A.C. Labonte, R. Robl, S. Shrotri, A.C. Grammer, and P.E. Lipsky. (2021, Mar.). "Comprehensive transcriptomic analysis of COVID-19 blood, lung, and airway." Scientific Reports. [On-line]. 11(1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8007747/pdf/41598_2021_Article_86002.pdf [June 25, 2021].

[31] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles." Proceedings of the National Academy of Sciences of the United States of America, vol. 102(43), pp. 15545-15550, Oct. 2005.

[32] Gene Ontology Consortium. (2021, Jan.). "The Gene Ontology resource: enriching a GOld mine." Nucleic Acids Res 2021. [On-line]. 49(D1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7779012/pdf/gkaa1113.pdf [June 25, 2021].

[33] M. Kanehisa, M. Furumichi, Y. Sato, M. Ishiguro-Watanabe, and M. Tanabe. (2021, Jan.). "KEGG: integrating viruses and cellular organisms." Nucleic Acids Research. [On-line].

49(D1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7779016/pdf/gkaa970.pdf [June 25, 2021].

[34] A. Liberzon, C. Birger, H. Thorvaldsdottir, M. Ghandi, J.P. Mesirov, and P. Tamayo. “The Molecular Signatures Database (MSigDB) hallmark gene set collection.” Cell Systems, vol. 1(6), pp. 417-425, Dec. 2015.

[35] A. Park and L.K. Harris. “Gene Expression Meta-Analysis Reveals Interferon-induced Genes Associated with SARS Infection in Lungs.” Frontiers in Immunology. [On-line] 12. Available: https://www.frontiersin.org/articles/10.3389/fimmu.2021.694355/full [June 25, 2021].

[36] E. Clough and T. Barrett. “The Gene Expression Omnibus Database.” Methods in Molecular Biology, vol. 1418, pp. 93-110, 2016.

[37] M.M. Becker, R.L. Graham, E.F. Donaldson, B. Rockx, A.C. Sims, T. Sheahan, R.J. Pickles, D. Corti, R.E. Johnston, R.S. Baric, and M.R. Denison. “Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice.” Proceedings of the National Academy of Sciences of the United States of America, vol. 105(50), pp. 19944-19949, Dec. 2008.

[38] T. Yoshikawa, T.E. Hill, N. Yoshikawa, V.L. Popov, C.L. Galindo, H.R. Garner, C.J. Peters, and C.T. Tseng. (2010, Jan.). “Dynamic innate immune responses of human bronchial epithelial cells to severe acute respiratory syndrome-associated coronavirus infection.” PLoS One. [On-line]. 5(1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2806919/pdf/pone.0008729.pdf [June 25, 2021].

[39] H.D. Mitchell, A.J. Eisfeld, A.C. Sims, J.E. McDermott, M.M. Matzke, B.J. Webb-Robertson, S.C. Tilton, N. Tchitchek, L. Josset, C. Li, A.L. Ellis, J.H. Chang, R.A. Heegel, M.L. Luna, A.A. Schepmoes, A.K. Shukla, T.O. Metz, G. Neumann, A.G. Benecke, R.D. Smith, R.S. Baric, Y. Kawaoka, M.G. Katze, and K.M. Waters. (2013, Jul.). “A network integration approach to predict conserved regulators related to pathogenicity of influenza and SARS-CoV respiratory viruses.” PLoS One. [On-line]. 8(7). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3723910/pdf/pone.0069374.pdf [June 25, 2021].

[40] H. Katsura, V. Sontake, A. Tata, Y. Kobayashi, C.E. Edwards, B.E. Heaton, A. Konkimalla, T. Asakura, Y. Mikami, E.J. Fritch, P.J. Lee, N.S. Heaton, R.C. Boucher, S.H. Randell, R.S. Baric, and P.R. Tata. “Human Lung Stem Cell-Based Alveolospheres Provide Insights into SARS-CoV-2-Mediated Interferon Responses and Pneumocyte Dysfunction." Cell Stem Cell, vol. 27(6), pp. 890-904 e898, Dec. 2020.

[41] Addinsoft. “XLSTAT statistical and data analysis solution.” 2019. Internet: https://www.xlstat.com/en/, 2021 [June 25, 2021].

[42] T.S. Fung and D.X. Liu. “Human Coronavirus: Host-Pathogen Interaction.” Annual Review of Microbiology, vol. 73, pp. 529-557, Sept. 2019.

[43] Z. Wehbe, S. Hammoud, N. Soudani, H. Zaraket, A. El-Yazbi, and A.H. Eid. (2020 Jun.). “Molecular Insights Into SARS COV-2 Interaction With Cardiovascular Disease: Role of RAAS and MAPK Signaling.” Frontiers in Pharmacology. [On-line] 11(836). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7283382/pdf/fphar-11-00836.pdf [June 25, 2021].

[44] S. Sugawara, D.L. Thomas, and A. Balagopal. “HIV-1 Infection and Type 1 Interferon: Navigating Through Uncertain Waters.” AIDS Research and Human Retroviruses, vol. 35(1), pp. 25-32, Jan. 2019.

[45] F. Lin and K. Shen. “Type I interferon: From innate response to treatment for COVID-19.” Pediatric Investigation, vol. 4(4), pp. 275-280, Dec. 2020.

[46] Y. Tang, J. Liu, D. Zhang, Z. Xu, J. Ji, and C. Wen. (2020, Jul.). “Cytokine Storm in COVID-19: The Current Evidence and Treatment Strategies.” Frontiers in Immunology. [On-line] 11. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7365923/pdf/fimmu-11-01708.pdf [June 25, 2021].

[47] V.J. Costela-Ruiz, R. Illescas-Montes, J.M. Puerta-Puert, C. Ruiz, and L. Melguizo-Rodriguez. “SARS-CoV-2 infection: The role of cytokines in COVID-19 disease.” Cytokine Growth Factor Reviews, vol. 54, pp. 62-75, Aug. 2020.

[48] Z. Chen, J. Hu, L. Liu, Y. Zhang, D. Liu, M. Xiong, Y. Zhao, K. Chen, and Y.M. Wang. “Clinical Characteristics of Patients with Severe and Critical COVID-19 in Wuhan: A Single-Center, Retrospective Study.” Infectious Diseases and Therapy, vol. 10(1), pp. 421-438, Mar. 2021.

[49] N. Peiffer-Smadja and Y. Yazdanpanah. “Nebulised interferon beta-1a for patients with COVID-19.” The Lancet Respiratory Medicine, vol. 9(2), pp. 122-123, Feb. 2021.

[50] A.B. Rowaiye, O.A. Okpalefe, O. Onuh Adejoke, J.O. Ogidigo, O. Hannah Oladipo, A.C. Ogu, A.N. Oli, S. Olofinase, O. Onyekwere, A. Rabiu Abubakar, . “Attenuating the Effects of Novel COVID-19 (SARS-CoV-2) Infection-Induced Cytokine Storm and the Implications.” Journal of Inflammation Research, vol. 14, pp. 1487-1510, Apr. 2021.

[51] J.S. Kim, J.Y. Lee, J.W. Yang, K.H. Lee, M. Effenberger, W. Szpirt, A. Kronbichler, and J.I. Shin. “Immunopathogenesis and treatment of cytokine storm in COVID-19.” Theranostics, vol. 11(1), pp. 316-329, Jan. 2021.

[52] C.M. Lloyd and R.J. Snelgrove. (2018 Jul.). “Type 2 immunity: Expanding our view.” Science Immunology. [On-line] 3(25). Available: https://immunology.sciencemag.org/content/3/25/eaat1604 [June 25, 2021].

[53] H. Ledford. “How 'killer' T cells could boost COVID immunity in face of new variants.” Nature, vol. 590(7846), pp. 374-375, Feb. 2021.

[54] T. Sekine, A. Perez-Potti, O. Rivera-Ballesteros, K. Stralin, J.B. Gorin, A. Olsson, S. Llewellyn-Lacey, H. Kamal, G. Bogdanovic, S. Muschiol, D.J. Wullimann , T. Kammann, J. Emgård, T. Parrot, E. Folkesson, O. Rooyackers, L.I. Eriksson, J.I. Henter, A. Sönnerborg, T. Allander, J. Albert, M. Nielsen, J. Klingström, S. Gredmark-Russ, N.K. Björkström, J.K. Sandberg, D.A. Price, H.G. Ljunggren, S. Aleman, and M. Buggert. “Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19.” Cell, vol. 183(1), pp. 158-168 e114, Oct. 2020.

[55] R.S. Abraham, J.M. Marshall, H.S. Kuehn, C.M. Rueda, A. Gibbs, W. Guider, C. Stewart, S.D. Rosenzweig, H. Wang, S. Jean, M. Peeples, T. King, W.G. Hunt, J.R. Honegger, O. Ramilo, P.J Mustillo, A. Mejias, M.I. Ardura, and M. Shimamura. “Severe SARS-CoV-2 disease in the context of a NF-kappaB2 loss-of-function pathogenic variant.” J Allergy Clin Immunol 2021, 147(2):532-544 e531.

[56] M. Kandasamy. “NF-kappaB signalling as a pharmacological target in COVID-19: potential roles for IKKbeta inhibitors.” Naunyn-Schmiedeberg's Archives of Pharmacology, vol. 394(3), pp. 561-567, Jan. 2021.

[57] J. Huang, A.J. Hume, K.M. Abo, R.B. Werder, C. Villacorta-Martin, K.D. Alysandratos, M.L. Beermann, C. Simone-Roach, J. Lindstrom-Vautrin, J. Olejnik, E.L. Suder, E. Bullitt, A. Hinds, A. Sharma, M. Bosmann, R. Wang, F. Hawkins, E.J. Burks, M. Saeed, A.A. Wilson,

E. Mühlberger, and D.N. Kotton. “SARS-CoV-2 Infection of Pluripotent Stem Cell-Derived Human Lung Alveolar Type 2 Cells Elicits a Rapid Epithelial-Intrinsic Inflammatory Response.” Cell Stem Cell, vol. 27(6), pp. 962-973 e967, Dec. 2020.

[58] E. Mick, J. Kamm, A.O. Pisco, K. Ratnasiri, J.M. Babik, C.S. Calfee, G. Castaneda, J.L. DeRisi, A.M. Detweiler, S. Hao, K.N. Kangelaris, G.R. Kumar, L.M. Li, S.A. Mann, N. Neff, P.A. Prasad, P.H. Serpa, S.J. Shah, N. Spottiswoode, M. Tan, C.S. Calfee, S.A. Christenson, A. Kistler, and C. Langelier. (2020, Nov.). “Upper airway gene expression reveals suppressed immune responses to SARS-CoV-2 compared with other respiratory viruses.” Nature Communications. [On-line] 11(1). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7673985/pdf/41467_2020_Article_19587.pdf [June 25, 2021].

[59] .B.K. Manne, F. Denorme, E.A. Middleton, I. Portier, J.W. Rowley, C. Stubben, A.C. Petrey, N.D. Tolley, L. Guo, M. Cody, A.S. Weyrich, C.C. Yost, M.T. Rondina, and R.A. Campbell. “Platelet gene expression and function in patients with COVID-19.” Blood, vol. 136(11), pp. 1317-1329, Sept. 2020.

[60] M. Wu, Y. Chen, H. Xia, C. Wang, C.Y. Tan, X. Cai, Y. Liu, F. Ji, P. Xiong, R. Liu, Y. Guan, Y. Duan, D. Kuang, S. Xu, H. Cai, Q. Xia, D. Yang, M.W. Wang, I.M. Chiu, C. Cheng, P.P. Ahern, L. Liu, G. Wang, N.K. Surana, T. Xia, and D.L. Kasper. “Transcriptional and proteomic insights into the host response in fatal COVID-19 cases.” Proceedings of the National Academy of Sciences of the United States of America, vol. 117(45), pp. 28336-28343, Nov. 2020.