University of Southampton

Faculty of Medicine

Clinical and Experimental Sciences

Airway epithelial cell regulation of T cell responses during viral infection

by

Claire Simms

Submitted for the degree of Master of Philosophy

October 2023

Abstract

The majority of acute respiratory tract infections are caused by viruses, with Influenza A viruses (IAV), Respiratory Syncytial virus (RSV) and Coronaviruses (CoV) three important causes of disease. The symptoms of infection can be very similar despite different aetiology and mechanisms of infection. Each virus can also result in a wide range of patient outcomes from asymptomatic infections to life threatening lower respiratory infections such as pneumonia. The ongoing COVID-19 pandemic is an important example of the considerable variation in patient outcome that can occur in response to a single viral agent, SARS-CoV2. Differences in how the host immune system responds to infection can contribute to this variation with the most severe cases frequently associated with excessive cytokine production and T cell dysregulation.

Given that T cells are a key factor in both viral clearance and immune mediated tissue damage, it is important to understand how different respiratory viruses impact T cell driven immunity. Epithelial cells that line the airways are often the primary target of respiratory viruses. These cells are increasingly understood to play a direct role in antiviral immunity, regulating and directing aspects of both the innate and adaptive responses at the site of infection. Signals released from epithelial cells can influence all stages of the T cell life cycle; both indirectly by influencing Dendritic cell polarisation and directly through T cell co-receptor expression as well as cytokine and chemokine production.

The primary aim of this work was to use transcriptomics data to compare T cell modulatory signals produced by respiratory epithelial cells in response to infection with different viruses. Published microarray data from experiments using well differentiated primary bronchial epithelial cells as models of RSV, IAV and SARS-CoV1 infection were re-analysed using differential expression analysis and gene expression was compared across viruses with a focus on T cell recruitment, activation and inhibition.

In addition, this project worked to develop an Air liquid interface model of RSV infection using the BCI-NS1.1 cell line with a view to validating the transcriptomics analysis. The validity of the model

was explored using RT-PCR to monitor RSV amplification over time as well as *CXCL10* expression as a marker of antiviral gene expression.

Considerable differences were observed in the transcriptomic responses between viruses. RSV and SARS-CoV1 upregulated far fewer genes than two strains of IAV. Focused exploration of T cell related genes identified differences in expression of chemokine genes as well as the cytokines *IL6*, *IL1B* and *IL23a*. The BCi-NS1.1 cell line showed promise as a candidate for an RSV ALI infection model, however consistent and reliable infection was not demonstrated despite more productive infection occurring in primary bronchial epithelial cells. While this suggests that differentiated BCi-NS1.1 cells are not permissive to infection, further work discussed in this report is needed to improve the model before this conclusion can be made.

Table of Contents

Table of Figures

Abbreviations

Abrv.	Definition
ALI	Air Liquid Interface
APC	Antigen Presenting Cell
ARDS	Acute Respiratory Distress Syndrome
ARTI	Acute Respiratory Tract Infection
B2M	Beta-2 Microglobulin
BSA	Bovine Serum Albumin
CC	Cellular component
CCL	Chemokine (C-C motif) Ligand
CD	Cluster of Differentiation
cDNA	Complementary Deoxyribonucleic Acid
COPD	Chronic Obstructive Pulmonary Disease
CoV	Coronavirus
COVID19	Coronavirus Disease 2019
Ct	Cycle Threshold
CXCL	Chemokine C-X-C motif ligand
DAPI	4',6-Diamidino-2-Phenylindole
DC	Dendritic Cell
DE	Differentially Expressed/ Differential Expression
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
dsRNA	Double Stranded Ribonucleic Acid
ECM	Extracellular Matrix
ELISA	Enzyme-Linked Immunosorbent Assay
FC	Fold Change
FDR	False Discovery Rate
GO	Gene Ontology
HA	Hemagglutinin
HAE	Human Airway Epithelial cell (Culture model)
IAV	Influenza A Virus
ICAM-1	Intercellular Adhesion Molecule 1
IFN	Interferon
IL	Interleukin
IRF	Interferon Regulatory Factor
ISG	Interferon Stimulated Gene
JAK	Janus Kinase
LRTI	Lower Respiratory Tract Infection
MERS	Middle East Respiratory Syndrome
MOI	Multiplicity Of Infection
mRNA	Messenger RNA
NF-κβ	Nuclear Factor kappa-light-chain-enhancer of Activated B Cells

ORA	Over Representation Analysis
PAMP	Pathogen Associated Molecular Pattern
PBEC	Primary Bronchial Epithelial Cell
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PFU	Plaque Forming Units
PRR	Pattern Recognition Receptor
QC	Quality Control
RLR	retinoic Acid-inducible Gene-I-Like Receptor
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RSV	Respiratory Syncytial Virus
SARS	Severe Acute Respiratory Syndrome
SD	Standard Deviation
ssRNA	Single Stranded Ribonucleic Acid
STAT	Signal Transducer and Activator of Transcription
TCA	The Citric Acid Cycle
Th	T Helper
TLR	Toll Like Receptor
TNF	Tumour Necrosis Factor
Treg	T Regulatory
Trm	Tissue Resident Memory
vRTI	Viral respiratory tract infections
WHO	World Health Organization

Chapter 1 Introduction

1.1 Respiratory viral infection

1.1.1 Overview

Viral respiratory tract infections (vRTI) are among the most common and important causes of disease worldwide, representing a significant contributor to infant mortality as well critical exacerbations of asthma or COPD (1,2). The clinical manifestation of vRTI can vary considerably from mild upper respiratory disease such as the 'common cold', to life threatening pneumonia, bronchiolitis and acute respiratory distress syndrome (ARDS) with severe lower airway involvement (3). Influenza viruses, particularly Influenza A virus (IAV), and Respiratory Syncytial Virus (RSV) are two of the most import causes of vRTI with IAV alone thought to be responsible for 3 to 5 million severe illnesses and approximately 290,000 to 650,000 deaths each year (4). Similarly, RSV is estimated to result in approximately 3.6 million hospitalisation due to lower respiratory tract infections (LRTI) and is the most important pathogen associated with LRTI in infants and young children (5–8). Due to the high prevalence of RSV, almost every child will have encountered RSV by the age of 2 with an estimated 2-3% of cases developing severe LRTI (8). Both RSV and IAV circulate simultaneously and in temperate climates produce seasonal peaks of disease which place considerable strain on healthcare systems (9).

The ongoing Coronavirus Disease 2019 (COVID-19) pandemic, caused by Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), highlights a further threat posed by novel strains of respiratory viruses. Since it emerged in China in late 2019, SARS-CoV2 has spread around the world resulting in over 765 million confirmed cases and over 6.9 million deaths, as of 7 May 2023 (10). SARS-CoV2 is the third coronavirus (CoV) to cause serious disease in the past century following from Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV-1) in 2003 and Middle East Respiratory Syndrome Coronavirus (MERS-CoV), present since 2012. SARS resulted in 8422 cases and 91 fatalities

before it was contained while MERS there were 2502 laboratory-confirmed cases and 861 associated deaths, primarily in Saudi Arabia (11,12). A key factor in these differences is the enhanced transmission efficiency of SARS-CoV2. Importantly, while each pandemic coronavirus (CoV) strain causes LRTI, SARS-CoV2 also replicates in the upper airways. This considerably impacts efforts to contain the virus as transmission can occur before the onset of symptoms or in asymptomatic patients (13,14).

Endemic strains of CoV make up a large proportion of seasonal cold viruses and produce mild or subclinical symptoms in the majority of individuals. In contrast to endemic viruses which co-evolved with the human population, pandemic viruses evolve from animal viruses (zoonotic) to which the human population bears no pre-existing immunity (15). Novel Influenza viruses have been the cause of 14 pandemics in the previous 500 years, most recently, the 2009 H1N1 ("swine Flu") pandemic caused an estimated 151,700 - 575,400 casualties globally (16). This virus is an evolutionary descendent of early swine influenza viruses and the 1918 H1N1 "Spanish flu" virus which killed up to 50 million people, the most lethal pandemic in recorded history. Subsequent outbreaks of a new H2N2 subtype in 1957 and H3N2 in 1968 resulted in more than 1 million and around 1 – 4 million deaths around the world respectively (17,18). Variants of H1N1 and H3N2 continue to circulate as endemic influenza viruses causing regular outbreaks as part of the seasonal pool of respiratory viruses (19). SARS-CoV2 variants are expected to transition to long-term persistence in communities possibly with the seasonal periodicity seen for influenza viruses (20).

Treatment for vRTI is primarily supportive and symptomatic (21,22). Uncomplicated infections often resolve within 1-2 weeks with over the counter (OTC) analgesics and decongestants used to ease symptoms as needed. The few antiviral drugs available are generally restricted for use in the most high-risk patients. An inhaled version of the antiviral drug Ribavirin is approved for treatment of severe RSV bronchiolitis in infants however multiple trials were not able to demonstrate clinically relevant benefits (22). Prophylactic administration of Palivizumab, the monoclonal anti-RSV-F

protein, confers some protection against severe disease, however the cost limits its use to children at the highest risk i.e.preterm babies or infants with congenital heart disease.

No specific RSV vaccines are currently in use and previous efforts have been unsuccessful, in some studies vaccination produced worse outcomes on reinfection. Although global efforts produced several vaccines against SARS-CoV2, the emergence of new variants reduced the efficacy of vaccine induced immunity to SARS-CoV2, necessitating repeat vaccination. Similarly, while influenza vaccines show some efficacy at reducing the severity of disease in at risk populations, vaccination strategies need to be revised each year in response to the high variability of circulating strains of IAV.

The wide variation in clinical manifestation and patient outcome is driven largely by the host immune response to infection (23–25). While an effective immune response is necessary for viral clearance and recovery, uncontrolled inflammation and immunopathology can be fatal. Significant inflammation resulting from excessive production of cytokines and chemokines ("Cytokine Storm") as well as disruption of T cell immunity frequently accompanies severe decline in respiratory viral infections (26). This presents a major challenge for antiviral drug and vaccine development and is a vital area of ongoing research.

1.1.2 Respiratory Virus Structure and Function

The structural characteristics of respiratory viruses inform aspects of their virulence including transmission rate and their ability to evade immune responses. Importantly, differences in antigenic receptor expression determine the host range of a virus while virion shape and genetic composition influence how well a virus can survive outside of a host or produce new variants.

RNA viruses, including IAV, RSV and CoV, are prone to mutations due to the error prone nature of RNA viral replication. This is seen in the high variability of IAV antigenic receptor expression which drive in the need to develop new influenza vaccines annually (15). Similarly, variations in SARS-CoV2

receptors have also led to the emergence of variants with a greater rate of infectivity and spread (27). The following section describes important structural and functional characteristic of RSV, IAV and SARS-CoV2 which contribute to virus and strain specific differences in infection.

RSV

RSV is an enveloped virus of the *Pneumoviridae* family which is classified into two subgroups, RSV-A and RSV-B. These antigenic subtypes circulate simultaneously to lesser or greater extents in different populations with RSV-A occasionally found to cause marginally more severe disease (28–30). The RSV virion contains a non-segmented, negative sense single strand RNA (− ssRNA) genome which is tightly associated with nucleoprotein (N) molecules to form the viral ribo-nucleoprotein complex (vRNP). Additional proteins required for polymerase function, phosphoprotein polymerase co-factor (P) and large polymerase subunit (L), are packaged alongside the RNA in the helical structured vRNP or nucleocapsid. A layer of matrix proteins (M) supports the outer envelope surrounding the nucleocapsid and give the virion its shape with both spherical and filamentous forms known to exist. Numerous copies of three main transmembrane glycoproteins, the attachment glycoprotein (G), fusion protein (F) and small hydrophobic protein (SH), project from the viral envelope. In addition to these structural proteins, the RSV genome encodes two non-structural proteins (NS-1, NS-2) involved in viral virulence and two regulatory proteins (M2-1 and M2-2) which are associated with transcription and replication (31–33).

The RSV-G protein defines the antigenic subtype of RSV and facilitates attachment to host cells via several proposed receptors. The specific attachment site may be cell type dependent, for example, binding of heparan sulphate proteoglycans (HSPGs) has been reported in immortalised cell lines but the CX3 chemokine receptor 1 (CX3CR1) may be of greater importance in primary human airway epithelial cells as well as *in vivo* ciliated epithelium, the preferential target of RSV (34).

After RSV is localised to the target cell, the F protein mediates fusion with the viral and host cell membranes. Fusion is catalysed by binding to host cell receptors such as nucleolin (35), this induces a conformational change in the F protein which both fuses the virion and host cell and forms the fusion pore. In addition to the plasma membrane, RSV fusion may also occur at the endosomal membrane following micropinocytosis (non-selective uptake into endocytic vesicles) (36).

Once the viral RNA is released into the cell, viral proteins associated with the nucleocapsid (N, P and L) initiate transcription and the resulting mRNAs are translated into viral proteins by the host machinery. Replication of viral genomes and assembly with viral proteins occurs in the cytoplasm of infected cells (31). New virions bud at or near the plasma membrane and infected cells subsequently express the G and F protein on their surface facilitating fusion and infection of neighbouring cells. This gives rise to the characteristic aggregates of infected cells (syncytia) seen during *in vitro* infection.

Influenza A

Influenza viruses belong to the *Orthomyoxoviridae* family and are classified into four main types (A-D) with influenza type A viruses (IAV) most likely to cause disease and pandemic outbreaks. IAV are categorised into strains based on the arrangement of hemagglutinin (HA) and neuraminidase (NA) proteins expressed on the host derived lipid envelope surrounding the complete virion (37). There are currently 18 HA and 11 NA variants known to exist in nature, most of which are found in avian and bat IAV reservoirs (37). The H1N1 and H3N2 IAV subtypes currently cause most epidemic disease in humans.

In contrast to RSV and other RNA viruses, the IAV genome is comprised of eight distinct segments of negative sense ssRNA. The segments are coated with viral nucleoproteins (NP) and combined with viral RNA-dependent RNA polymerase (RdRp) subunits (PA, PB1, PB2) into vRNP complexes (38). Whole sections of the IAV genome can mix and recombine with segments from other IAV viruses

during co-infection. If this occurs with HA or NA encoding segments, antigenic shift can occur resulting in a new virus with un-familiar antigenic properties. This is thought to have contributed to the emergence of the 1918 H1N1 and 1968 H3N2 pandemics which originated from re-arrangements which exchanged avian HA segments with mammalian IAV (15).

Viral entry is facilitated by HA proteins which bind terminal sialic acid (SA) residues on cell surface glycoproteins and glycolipids. Affinity of HA for SA residues contributes to cell tropism and inter-species transmission. For example, human seasonal IAV generally bind SAα-2,6-Gal type receptors found in the upper airways, however, the avian IAV receptor type SAα-2,3-Gal has been found in the lower airways of humans, raising the possibility of zoonotic infection (39). Internalisation may require activation of additional subtype dependent internalization receptors, such as nucleolin or EGFR to trigger endocytosis. The low pH inside the endosome induces a conformational change in HA which leads to fusion of the viral and endosomal membranes. Further low pH driven changes in the matrix (M) protein, which like RSV surround the vRNP beneath the viral envelope, contributes to un-coating of the virus and release of vRNPs. Transcription and replication takes place in the nucleus which is facilitated by nuclear localisation signals (NLS) expressed on NP proteins (40,41). Release of nascent virions is aided by NA glycoproteins which cleave SA moieties on the cell surface which might otherwise bind with HA. This also facilitates the dispersion of the new virion and is a key factor in the enhanced cell to cell spread when compared to RSV.

SARS-CoV2

SARS-CoV2 has been classified as a Beta-Coronavirus (BetaCoV) from the *Coronavirinae* subfamily of *Coronaviridae*. This classification also contains MERS-CoV and SARS-CoV as well as the less pathogenic human CoV strains HCoV-OC43 and HCoV-HKU1 (42). Based on sequence similarity, SARS-CoV2 was grouped into the *Sarbecovirus* sub-genera of BetaCoV alongside SARS-CoV1.

All coronaviruses contain a relatively large positive sense single strand RNA (+ssRNA) genome which, like RSV, is packaged into a helical shape with multiple nucleocapsid (N) proteins. The genome encodes four major structural proteins; the spike (S) protein, membrane (M) protein, N protein and envelope (E) protein, which are expressed to varying degrees in complete CoV virions (43)(44). The M protein is most abundant and is embedded in the outer lipid membrane where it gives the virion its shape and structure. Less if known about the E protein which is found in lower numbers and may enhance infectivity by forming pores in host cell membranes (43)(44).

The large S glycoprotein projects from the spherical CoV envelope in groups of three forming the crown or "corona" like appearance which gives CoV their name. This protein mediates binding and entry to host cells and may be an important factor in determining the virulence of different coronaviruses. Both SARS-CoV1 and SARS-CoV2 enter cells via the receptor *angiotensin-converting enzyme 2 (ACE2),* which is expressed on airway epithelial cells from the upper and lower airways as well as cells from the heart, kidneys and reproductive tracts (45). Variations in the S protein has resulted in SARS-CoV2 having 10-20 fold greater affinity for ACE2 than SARS-CoV1 (46). Mutations in the S protein have also been key factors in the emergence of recent variants of concern which show enhanced affinity for ACE2 receptors as well as reduced neutralisation ability of targeted antibody responses (27). Differences in ACE2 expression levels in the respiratory tract mirror the SARS-CoV-2 infection gradient, with highest expression in nasal ciliated cells, the primary targets for early SARS-CoV-2 replication.

In SARS-CoV2, the S protein is cleaved into functional S1 and S2 domains during maturation of the virion in infected cells. These domains remain as non-covalently associated subunits of the S protein; S1 binds ACE2 on host cells while S2 is anchored to the virus membrane and mediates membrane fusion (46). ACE2 engagement induces a conformation change which exposes S2 to additional cleavage by host proteases. The transmembrane protease serine 2 (TMPRSS2) is thought to be

important for this stage during entry at the host cell membrane. SARS-CoV2 may also enter cells via the endocytic pathway with S2 activation mediated by host cathepsins.

Cleavage of S2 releases a fusion peptide which initiates formation of a fusion pore allowing entry of the viral genome. Replication and translation of viral RNA then occurs in the cytoplasm with involvement of many of the 16 non-structural proteins (nsp1 – nsp16) encoded by the SARS-CoV2 genome. New virions are assembled in the cytoplasm and are secreted from the infected cell by exocytosis, in addition, new virions may pass directly to other cells through membrane fusion.

1.1.3 Pathology of Respiratory Viral infection

Respiratory viruses are primarily transmitted in respiratory secretions with contact occurring through direct contact with individuals, indirect contact with contaminated surfaces or through inhalation of respiratory droplets or aerosols. The rate of symptom onset and disease severity will depend on host risk factors such as age, obesity, or immunosuppression as well as prior immune memory.

In symptomatic influenza infections, patients experience a sudden onset of high fever, cough, headache and malaise which often persist for 7 to 10 days, this can be followed by fatigue and weakness which may linger for weeks (47). Pathology can be seen in the lung soon after symptom onset, for instance, in uncomplicated influenza, superficial inflammatory lesions are present in tracheal and bronchial biopsies as early as 1-6 days after the emergence of symptoms (48).

This symptom profile aligns with findings from experimental human IAV infection in which mean viral load and symptom severity peak on day 2 and 3 following inoculation respectively. In the same study, RSV replication was more delayed, with an average 4.6-day incubation period until first detection in respiratory secretions (49). Peak viral load occurred on day 5.4 with symptom severity increasing with viral load and like IAV, reaching peak scores 24h later.

In clinical infant infections, RSV can manifest 4-6 days after exposure with nasal congestion and a mild fever, followed by a cough and wheezing in cases of bronchiolitis. The histopathology of autopsy samples from infants who died from bronchiolitis show an otherwise normal appearance in most RSV positive cells with damage most often concentrated around the bronchi and bronchioles (50). This differs to severe infections with IAV in which increasing epithelial cell damage occurs due to both viral cytolysis and inflammatory cell infiltration (48).

Lower airway involvement in COVID-19 presents with a characteristic bilateral pneumonia with damage to the alveolar epithelium and degradation of the extracellular matrix (ECM) that maintains the structure of the lung (51). Involvement of significant portions of the airways affects lung function and may result in ARDS, in these cases the patient often requires mechanical ventilation. While most patients develop mild symptoms including a cough, low grade fever and occasionally loss of taste or smell, which resolve easily, over 15% of patients develop severe disease with mortality as high as 61.5% in critical cases (52).

An accumulation of immune cells in the airways is a common feature in autopsy studies of fatal SARS CoV-2 infection (51). Similarly, in RSV bronchiolitis, occlusion of the bronchiolar and alveolar lumen may be exacerbated by immune cell infiltration (50,53). It is hypothesised that the relatively mild tissue dame seen in RSV, in contrast to the extensive cytopathology observed in other severe LRTI, could be driven by immune cells (54). However, the factors which drive this remain to be understood.

While controlled production of inflammatory mediators helps to clear and infection, excessive or prolonged exposure can destroy tissues leading to ARDS and multi-organ failure. Pro-inflammatory cytokine production is initiated by infected airway epithelial cells to control the infection and recruit immune cells to the site. Cytokine levels increase further as infiltrating immune cells are activated and without resolution of infection, their continued production results in cytokine storm and deterioration of the patient.

1.2 The Respiratory Epithelium

The morphologically distinct cells that make up the respiratory epithelium are the primary entry point for respiratory viruses and are a central component of the immune response in the lung (55). The upper airways (nasal cavity and pharynx) and conducting portion of the lower airways (trachea, bronchi and bronchioles) are comprised mainly of pseudostratified ciliated epithelial cells and mucus secreting goblet cells above a layer of replicative basal cells lining the basement membrane (56). Tight cell-cell junctions in the epithelial layer form an initial mechanical barrier to infection, blocking access to many viral entry receptors. Mucins produced by goblet cells contribute to a mucus layer which coats the airway lumen and traps pathogens and debris. Co-ordinated cilia movements direct this mucus layer towards the upper airways where it can be removed, protecting the lower airways from infection. Impairment of muco-ciliary clearance due to loss of cilia is an important factor in airway disease such as COPD or in smokers, increasing susceptibility to LRTI in these patients (57). Airway epithelial cells also release numerous enzymes and antimicrobial molecules such as nitric oxide, B-defensins and surfactant proteins which bind numerous viruses, including IAV, RSV and possibly SARS-CoV2, enhancing their clearance from mucosal points of entry (58).

In the smaller diameter bronchioles, the epithelial layer takes on a simple cuboidal structure and contains fewer basal cells and goblet cells than the more proximal airways (56). The terminal bronchioles become dominated by non-ciliated Club cells which have basal cell like replicative properties and a secretory function, producing antimicrobial substances including surfactants and uteroglobin (aka clara cell secretory protein). The alveolar compartment is lined with large membranous pneumocytes, known as Type I alveolar epithelial cells (AEC I) and granular pneumocytes known as Type II alveolar epithelial cells (AEC-II) (59). AEC-I are flattened wide squamous cells which cover approximately 90-95% of the alveolar surface area. Their thin cytoplasm enables gas exchange with microvascular endothelial cells that surround the alveoli. AEC-II are

smaller cuboidal cells which make up around 7% of the alveolar surface. They are responsible for regeneration of damaged AEC-I and contribute to alveolar defences and ion transport. AEC-II also produce surfactants which are essential for lowering surface tension in the alveoli and can be antimicrobial (60). In addition to ciliated cells in the larger airways, AEC I and AEC II cells are primary targets for respiratory viruses (50,61).

1.2.1 Viral Sensing and the Early Antiviral Response

As the portal for entry during respiratory viral infections, airway epithelial cells have a central role in initiating the immune response when viruses breach intrinsic defences. Epithelial cells express pathogen recognition receptors (PRR) including Toll like receptors (TLRs) and retinoic acid inducible gene-I (RIG-I)-like receptors (RLRs) which initiate antiviral signalling cascades in response to activation by foreign antigens known as pathogen associated molecular patterns (PAMPS) (62). PRRs respond to both extracellular and intracellular PAMPS based on their cellular location. For example, TLR2 and TLR4 are expressed on cell membranes and are activated by viral glycoproteins such as IAV-HA (63,64). TLR3 responds to dsRNA formed as an intermediary of viral replication and is located both within endosomal membranes and on the cell surface while RLRs including RIG-1 and MDA5, are activated by ssRNA in the cytosol (65,66). Activation of these receptors initiates signalling pathways that converge on the activation and translocation of interferon regulatory factors (IRF), importantly IRF3 and IRF7, and nuclear factor κβ (NF-κβ), which ultimately results in transcription of pro-inflammatory cytokines and generation of an antiviral state at the site of infection (67).

Production of both Type I (IFNα and IFNβ) and type-III interferon (IFN-λ) is a central consequence of PRR sensing in response to viral infection. Both classes result in activation of similar antiviral pathways with differences primarily in receptor expression and abundance (68). Almost all nucleated cells secrete IFNα/β and the type I IFN receptors (IFNAR1 and IFNAR2) are ubiquitously expressed. Type III IFN responses are more restricted to mucosal surfaces prone to pathogen exposure and IFN-

λ as well as the type III IFN receptors (IFNLR1 and IL-10Rβ) are preferentially expressed by epithelial cells on mucosal surfaces (62). Activation of either IFN receptor triggers signal transducer and activator of transcription 1 (STAT1) and STAT2 signalling cascades which lead to rapid transcription of hundreds of interferon-stimulated genes (ISGs) with broad antiviral functionality (69).

Cellular expression of the IFN-induced transmembrane protein (IFITM) family member IFITM3 for example, has been shown to impair early RSV entry and replication (70). Additional cytokines as well as chemokines which regulate the inflammatory response and immune cell recruitment, are simultaneously upregulated and released by epithelial cells alongside IFNs in response to PRR engagement and IFNR activation. This can occur rapidly, for example NF-κβ and IRF-3 activation induces expression of IFN-β, interferon inducible protein-10 (IP-10)/CXCL10, regulated and normal T-cell expressed and secreted (RANTES)/CCL5, and ISG-15 early into RSV infection (71). Similarly, there is rapid upregulation of IL-6, TNF-α, IL-8/CXCL8, CXCL10, CCL2, and CCL5 in response to infection with IAV (72). Acting locally, these cytokines/chemokines facilitate antiviral responses and recruit immune cells to the site of infection.

Differences in these markers often correlates with disease severity, In RSV infection, infants with severe bronchiolitis had lower nasal viral load and reduced CXCL10, CCL5 and IFN-γ levels compared to mild or moderate infection (73). Conversely, another study reported highest viral load and increased IL-6 levels in hospitalized older adults with RSV disease when compared to outpatients, further highlighting the importance of the inflammatory response to the varied outcome seen in different populations (7).

1.3 The Role of T cells in Respiratory Viral Infection

A key consequence of cytokine and chemokine signalling during infection is recruitment and activation of T cells; specialised lymphocytes with varied roles in antiviral immunity (74). Defects in T

cell immunity due to aging, immune suppression, or cancer are associated with more severe infections and poorer outcomes (75). In contrast to the immediate IFN driven innate response, adaptive T cell responses take days to weeks to be established and are targeted to an invading pathogen due to diverse antigen-specific T cell receptor (TCR) expression.

Broadly classified into CD8+ and CD4+ subtypes, T cells undertake a range of effector functions including clearance of infected cells, regulation of the inflammatory response and long term immune memory (74). CD8+ cytotoxic T lymphocytes (CTL) are the main drivers of viral clearance in the adaptive response. They produce antiviral cytokines including TNF-α and IFN-γ as well as cytotoxic granules which perforate cell membranes and induce apoptosis of infected cells. This targeted response is initiated following interaction between the TCR and target antigen (Ag) displayed on major histocompatibility complex molecules (MHC-I), expressed at the surface of nucleated host cells (76). High diversity in the binding domains of MHC-I molecules allows for presentation of a large repertoire of peptides to CD8+ T cells. Endogenous proteins are processed to MHC-I using the direct presentation pathway. During infection, MHC-I expression is upregulated from baseline low levels in the lung epithelium and protein fragments derived from viral replication are processed for presentation to CD8+ T cells alongside other cytosolic peptides.

Viral Ag can also be bound to MHC-I from exogenous proteins using the cross-presentation pathway (CP) (76). By this pathway, external Ag such as phagocytosed virus or infected cell fragments, can be acquired by professional Ag presenting cells (APC), such as Dendritic cells (DC), at the site of infection. This is important for the initial priming and expansion of Ag specific naïve T cells in the peripheral lymph nodes.

Cross presentation of external proteins to T cells is primarily associated with class 2 MHC molecules (MHC-II) expressed by APC. MHC-II are recognised by a highly diverse range of CD4+ T cells which play a major role in supporting CTL effector functions in addition to orchestrating other aspects of adaptive immunity (74). APC mediated differentiation and expansion of naïve CD4+ T cells occurs in

the lymph nodes following TCR recognition of cognate Ag in complex with MHC-II. Based on costimulatory signals and cytokine profiles during activation, several distinct CD4+ T helper (Th) cell lineages are produced with varied roles in antiviral immunity (77).

Perhaps most relevant to viral infection, activated Th1 type CD4+ T cells are characterised by production of IFN-y, TNF-α, and IL-2 which support the initial expansion of CTL as well as their effector activity. Meanwhile, Th2 cells produce cytokines such as IL-4 and IL-5 which downregulate Th1 functions and support activation of B cells, mast cells and basophils in favour of a humoral or anti-parasitic response. Similarly opposing roles have been prosed for the Th17 and Treg lineages (78). Th17 cells, characterised by IL-17 expression, may upregulate proinflammatory cytokine production and CTL activity. In contrast, Treg cells produce IL-10 and TGF-B and may suppress the effector functions of cytotoxic cells to limit immunopathology and viral induced lung damage.

A fraction of effector CD4+ and CD8+ T cells persist in the lung as antigen specific memory cells after disease resolution (79). These cells can remain in the lung long term and are capable of rapidly responding to repeat infection without reliance on additional co-stimulatory signals. High numbers of pre-existing antigen specific CD8+ T cells were found to correlate with less severe illness after infection with pandemic H1N1 (80). Similarly, in experimental human IAV infection, individuals with pre-existing CD4+ T cells specific for internal viral proteins, developed less severe symptoms, in the absence of humoral immunity (81). Memory Treg populations have been shown to rapidly respond to secondary infection and provide balance between viral clearance and immunopathology, suppressing CD4 and CD8+ T cell trafficking, activation and cytokine production (82). The ability to mount a response against conserved internal proteins means that T cell immunity can provide protection against multiple viral subtypes. For example, pre-existing SARS-CoV1 specific T cells have shown cross reactivity to SARS-CoV2 antigens (83). These memory T cell responses may also last longer than antibody driven immunity as has been seen in COVID-19 patients (84).

Finally, primed T cells can interact with each other and may attract other antigen specific clones during viral infection (85).This occurs through direct MHC-II interaction as well as cytokine secretion and can facilitate coordination of the wider immune response. Crucially, while not discussed here, CD4+ cells initiate the humoral response by activating B cells.

1.3.1 Dysregulation during viral infection

Activated CD8+ T cells appear in tracheal aspirates from patients with mild RSV, IAV and human CoV infections within the first 10 days of symptom onset (86). This rise in CTL numbers is associated with viral clearance, yet deficiencies in CD4 and cytotoxic CD8 T cells have even been associated with fatal influenza virus and RSV infections in infants (87).

RSV specific CTL numbers peak in the second week of symptom onset and gradually decline to absence by the eighth week. A murine RSV study has shown that CD8+ T cell recruitment to the lung may be co-ordinated by Tregs which may then reduce disease severity by limiting the strength of antigen specific CD8+ T cell responses and their production of TNF-α (72). Other studies have identified correlation between disease severity and expression of distinct populations of T cells. Patients with severe disease have shown Increased levels of CD8+ T cells expressing IL4+ as well as reduced proportions of CD8+ T cells expressing IFNγ (88). In this study, nasal aspirates from patients with severe RSV disease also had reduced concentrations of IL-17. Patients with greater frequencies of both CD8+ and CD4+ T cells expressing IL17 had shorter durations of hospitalization. Disruptions in TH1/ Th2 responses has been particularly implicated in severe RSV infections including previous poor response to vaccinations (89). There is also correlation between disease severity and a decline in peripheral T lymphocytes as well as the ratio of T cell subtypes with lower levels of regulatory T cells seen in more severe cases of COVID-19 (26). There is a trend towards decreasing Treg levels in the peripheral blood of COVID-19 patients (90). In addition, T cells from COVID-19 patients have

shown *in vitro* production of TNF, IFN-γ, IL-2 and IL-17 as well as skewing toward the Th17 phenotype (91).

Impaired CTL responses, in which virus specific CD8+ T cells fail to perform cytotoxic functions at the site of infection, have also been associated with disease progression and severity during IAV and SARS-CoV2 infection (92). In COVID-19, a weak and relatively delayed (> 15 days post symptom onset) induction of SARS-CoV2 specific T cells was seen only in patients with severe infections. In contrast, early detection was associated with accelerated viral clearance and a milder disease course.

1.3.2 T Cell Regulation by Respiratory Epithelial cells

T cell functionality needs to be controlled to minimise immunopathology while maximising viral clearance. This occurs though complex cytokine/chemokine driven cross talk between cells of the innate and adaptive immune response and the surrounding tissue. As early initiators of antiviral immunity, epithelial cells play a key role in mediating this communication and influencing a developing T cell response from activation to final memory cell formation. The importance of the epithelium-immune cell interaction in disease severity was recently demonstrated in a single cell sequencing analysis of bronchial samples from COVID-19 patients (93). In this study, chemokine and chemokine receptor expression were elevated in critical cases compared to moderate cases and ligand-receptor mapping demonstrated strong interactions between epithelial cells and immune cells, including CTLs, CD4+T/Treg.

The polarity of a developing T cell response depends to a large extent on the activation state of outgoing DCs and the resulting cytokine and co-stimulatory signals given to naïve T cells during their activation in the lung draining lymph nodes (94). While TLR signalling and antigen type are important factors, epithelial derived cytokines also influence DC activation state and subsequent T

cell expansion. Epithelial derived thymic stromal lymphopoietin (TSLP) has been shown to skew DCs so that they induce expansion of Th2 cells (95). Type I IFNs released from infected epithelial cells are likely to drive responses toward Th1 development, while Type III IFNs may skew responses toward a Treg driven state (96). The epithelial derived cytokines/chemokines CCL20 and Granulocyte-monocyte colony stimulating factor (GM-CSF) also support T cell activation by inducing DC mobilisation into the lung from the circulation and contributing to DC differentiation from monocytes respectively (97,98).

While DC mediated T cell activation in the lymph nodes is well established, there is also evidence to suggest T cell activation may occur directly in the lung. Mice lacking peripheral lymph nodes have been shown to generate antigen specific CD8+ T cell responses and clear low dose influenza infection (99). Similarly, CCR7−/− mice with impaired DC tracking from the lungs to the lymph nodes, were able to activate Mycobacterium tuberculosis specific CD4+ T cell proliferation in the lung following infection (100). Airway epithelial cell derived cytokines and chemokines may therefore also play a more direct role in DC and T cell activation.

Chemokines make up a significant proportion of the inflammatory mediators released by airway epithelial cells in responses to viral infection (71). These proteins are divided into four distinct groups, C, CC, CXC, and CX3C (based on NH_2-terminal cysteine residue number and spacing) that bind chemokine receptors expressed by distinct leukocyte subsets, inducing their migration to the site of infection. CXCR3 is rapidly induced on naïve cells following activation and remains highly expressed on Th1-type CD4+ T cells and effector CD8+ T cells (101). Airway epithelial cells enhance recruitment of these effector T cells to the lung through production of CXCR3 ligands, CXCL10, CXCL11 and CXCL9. Th2 type CD4+ T cells express CCR4 and CCR8 receptors that respond to the CCR4 ligand CCL22 and CCR8 ligand CCL1, both produced by airway epithelial cells in addition to the Th2 recruiting cytokine IL-1β (71,101). The expression of these chemokines is upregulated by cytokines released from recruited T cells in a feedback loop which amplifies the T cell polarity in the lung.

The pulmonary epithelium can also impact the overall inflammatory state in the lung by enhancing recruitment of other immune cells which promote different T cell lineages (102). Neutrophil infiltration for example increases Th1 and Th17 responses and eosinophil recruitment enhances Th2 responses. Epithelial cells are also a source of the IL-1 family member IL-33 which is a ligand for ST2, a receptor that strongly induces Th2 cytokine production (96).

Additionally, airway epithelial cells directly interact with T cells though expression of membrane-bound and soluble molecules. During active viral infection, epithelial cells will upregulate MHC-I expression to heighten antigen presentation to CTL. They also stimulate the secretion of transforming growth factor beta (TGF-β), which leads to the activation of regulatory T cells. In contrast, the inhibitory molecule PD-L1 (CD274) mediates T-cell apoptosis and functional exhaustion when engaged by its receptor PD-1 expressed by T cells. Persistence of PD-1+ CD8+ "exhausted" T cells with impaired cytotoxic functionality in the lung has been associated with disease progression and mortality in both COVID-19 and IAV (103)(92). These cells function normally at other sites such as the spleen and LDLN suggesting that T cell exhaustion is driven by factors within the lung. Inhibition of PDL-1 in a co-culture model of CD8+ T cells and bronchial epithelial cells infected with RSV, resulted in increased T cell effector functions, including IFN-γ, IL-2, and granzyme B expression and decreased viral load in the epithelial cells (104). PD-1/PD-L1 signalling may also promote formation of Trm cells and may be important for maintaining homeostasis in the healthy lung by inhibiting T cell function and cytokine secretion.

1.4 Aims and Hypothesis

T cell regulatory signals expressed by infected epithelial cells contribute to the nature and extent of the inflammatory response to viral infection. Given the importance of these responses to both viral clearance and immunopathology, understanding the complex communication between the host

epithelium and T cells and how they differ between viruses, is fundamental to identifying mechanisms of pathogenicity and potential therapeutic targets.

This project aimed to explore how three important respiratory viruses, (IAV, RSV and SARS-CoV1) influence bronchial epithelial cell expression of T cell regulatory genes using publicly available transcriptomics data. Comparing the transcriptomic responses of three RNA viruses, each with similar presentation and a varied patient outcome, will help to identify any differences in activated and suppressed T cell associated genes or pathways.

A secondary aim was to establish a reliable model of RSV infection using the newly created BCi-N1.1 cell line in Air-Liquid interface culture. Using a commercial cell line with ALI differentiation capabilities offers the possibility of increased consistency and better availability than primary bronchial epithelial cells (PBEC) obtained from patient samples. If successful, this model would be used to validate the PBEC-T cell gene signature generated from the transcriptomics analysis.

The primary hypothesis states that the expression of genes that impact T cell activation by virally infected bronchial epithelial cells will differ depending on the viral agent with which the cells are infected.

The specific aims of the project were to:

1. Perform differential expression analysis of virally infected bronchial epithelial cells in previously published transcriptomics datasets.
2. Compare expression of T cell modulatory genes between datasets for different viruses
3. Optimise an in vitro air-liquid interface (ALI) model of BCi-NS1.1 RSV infection

Chapter 2 Materials and Methods

2.1 Bioinformatics analysis

2.1.1 Data set Selection

Publicly available microarray datasets were obtained from the Gene Expression Omnibus (GEO) repository. Datasets were found by entering the terms “RSV” OR “Influenza” OR “SARS” AND “epithelial” into the GEO datasets advanced search builder. The results were filtered by species (*Homo sapiens*) and then manually searched to identify experiments that met the following criteria:

1. Primary human airway epithelial cells were used in an ALI culture model
2. Uninfected or mock infected samples were included
3. Each treatment group contained at least 3 replicate samples

The datasets chosen for this work are summarised in **Table 1.** Where additional treatment conditions are included in the full dataset but not used for this project, the chosen variables are highlighted bold. Citations refer to published research articles using the GEO data. Data sets can be accessed from the NCBI GEO database.

Table 1. Details of Microarray data sets used in this project

DataSet	Platform	Virus(es)	Model	Summary	Ref.
GSE32138	GPL6480 Agilent- Human Genome Microarray G4112F	**RSV244** MOI 5 **IAV** MOI 1 **(A/Udorn/307/1972/H3N2)**	ALI differentiated Primary Human bronchial epithelial cells	**Infection duration**: 2h apical **RNA Harvest HPI:** IAV (**24h**). RSV (**48h**). **Controls**: Time matched media only **Replicate arrays**: 4 per condition **Total Arrays**: 16/16	(105)
GSE47961	GPL6480 Agilent- Human Genome Microarray G4112F	**SARS-CoV1** MOI 2 **IAV** MOI 2 **(A/CA/04/2009/H1N1**) Not used: SARS-ddORF6 SARS-BatSRB	ALI differentiated Primary Human-tracheobronchial epithelial cells	**Infection Duration**:.2h apical **RNA Harvest HPI**: SARS-CoV1 (0, 24, 36, 48, 60, 72, 84 and 96 h). IAV (0, 6, 12, 18, 24, 36 and 48h) **Controls**: Time matched media only **Replicate Arrays:** 3-4 per condition using the same cell stock for all replicates **Total Arrays Used:** 88/141	(106)

Microarray data sets obtained from NCBI Gene expression omnibus (GEO) repository were chosen based on their use of ALI differentiated airway epithelial cells in respiratory viral infection experiments. Details of these data sets can be obtained from the NCBI GEO website.

2.1.2 Pre-Processing and Quality Control

Data set re-analysis was performed using RStudio (v 4.1.0) with the R programming language (107). Experiment metadata was obtained from the NCBI GEO database and imported into RStudio using the *GetGEO* function from the R package *GEOquerry* (108). This function also imports normalised expression data provided by the authors, however, to maintain consistency between datasets, raw data files were downloaded manually from supplementary files provided on the NCBI GEO website. As both data sets were generated using the same microarray platform, they were imported and pre-processed using the same overall pipeline of functions from the Linear Models for Microarray data (*limma*) package (109) (R script: Appendix A1 and A2). Unless otherwise stated, all functions described here are part of the *limma* package. Raw array intensity files were imported to RStudio using the *read.maimages* function with source specified as "Agilent" and single colour array indicated with green only = TRUE. Control probes and background intensities were included with the raw data, allowing for normalisation using the *neqc* function. By this method, the background probe intensities are first subtracted from foreground intensities and then the negative control probes are used for normal + exponential (normexp) background correction. An offset is added to the background-corrected data followed by quantile normalisation using both positive and negative control probes. The normalised intensity values are then log2 transformed and the control probes removed. After *neqc* normalisation, intensities from replicate spots were averaged and boxplots were generated to confirm successful normalisation.

The data were further explored using multidimensional scaling (MDS) to graphically represent the relationship between the arrays. The *plotMDS* function applies a variation of traditional MDS or principal components analysis particularly appropriate for microarray data. A Euclidian (root-mean-square deviation) distance measure is computed for each pair of samples in the data set, using the genes which most distinguish those samples. The default function uses different genes for each pairwise comparison, ensuring that distances between samples on the resulting scatterplot

approximate the leading log2-fold change between samples. Samples were coloured by group membership (infected/uninfected) or time point as required. Hierarchical clustering using wards linkage was also used to generate dendrogram's from the normalised data. Finally, probe annotations provided by the authors were updated using the *clusterProfiler* package (110) with the *hgug4112a*.db database and unlabelled probes were removed prior to differential expression (DE) analysis.

2.1.3 Differential Expression Analysis

DE analysis was performed for all data sets to compare expression between infected and mock-infected samples (R Script: Appendix A3). A standard limma model fit was applied with experimental variables defined in a design matrix. The chosen comparisons were then applied to the model as a contrasts matrix and the empirical Bayes (eBayes) procedure was used to determine test statistics including log2 fold change (log2FC) between stated comparisons, P values and false discovery rate (FDR) adjusted P values computed to control for multiple hypothesis testing. Limma uses the Benjamini and Hochberg's (BH) method to control for the FDR. In practice, this means that for a threshold FDR value of 0.05, the expected proportion of false discoveries should be less than 5%. For this analysis, genes were considered DE with FDR values below 0.05 and an effect size greater than log2FC 1 in either direction (log2FC >= 1 or <-1). The result of this analysis genome wide was visualised as volcano plots generated using the *EnhancedVolcano* package (111). DE gene counts were determined at the gene level by first averaging test statistics of genes with multiple probes using the *aggregate* function. The overlap in DE genes within each experimental group, were visualised using the *VennDiagram* package (112).

2.1.4 Time Series Analysis

Normalised (neqc), gene level expression data from the GSE47961 H1N1 and SARS-CoV1 arrays were additionally analysed as independent time series using the *maSigPro* package (113). This package applies a two-step regression strategy to identify significant changes in gene expression over time and between experimental groups. First, a global regression model is applied to each gene with regression variables (eg H1N1vsMOCK) defined in a design matrix. Genes with significant changes are selected with a threshold FDR of 0.05. Stepwise regression is then applied to compare differences between experimental groups and identify significant differences in expression profiles over the course of infection.

Cluster analysis using the "hclust" (Hierarchical clustering) method with "ward.D" agglomeration was used to group significant genes with similar expression patterns based on the correlation "cor" distance metric. The average intensity values for genes in each cluster are presented as graphs showing results from both infected and control arrays at all time points. R script for this analysis is included in Appendix A4.

2.1.5 Data Analysis: Pathway Annotation

Significantly DE genes from the GSE32138 *limma* analysis as well as important maSigPro clusters and significantly DE genes from peak time points from the GSE47971 analysis were annotated using the *clusterProfiler* R package. The enrichGO function was used to perform over representation analysis (ORA) for gene ontology Biological process (GO:BP) terms. Lists of upregulated and downregulated genes were analysed separately and were ordered by log2FC. Significant, but not log2FC filtered, genes were used as background or "universe" gene sets (R Script: Appendix A5).

Heatmaps were generated for genes with shared functional annotation, from gene level normalised intensity values, using the *coolmap* function from *limma*. In all cases, rows (genes) were scaled as Z-

Scores showing variation around the mean for a given gene. Both rows and columns (samples) were clustered using wards linkage for the GSE32138 data and columns were arranged by increasing time point for the GSE47961 data to allow easy visual comparison across the course of infection.

GO:BP terms relating to T cell recruitment or activation which were identified for each ORA, were reduced based on biological relevance to airway epithelial cell and T cell communication. The effect of each virus on DE of these genes were compared within datasets as both heatmaps of expression data and log2FC results.

2.2 In Vitro culture Models

2.2.1 Cell lines

Vero E6 cells are a well-established immortalised cell line derived from African green monkey kidney cells. There is a long history of publications using this cell line for both propagation and quantification of RSV. They were used here during initial infections due to their known permissivity to infection.

BCi-NS1.1 cells (114) (referred to as BCi throughout) are an immortalised human basal cell line derived from the large airways of a healthy non-smoker. They retain the characteristics of the original primary cells with multipotent differentiation capacity for over 40 passages. When grown under Air Liquid Interface (ALI) conditions, BCi cells form a pseudostratified layer comprised of ciliated columnar epithelium, mucus producing goblet cells and a basal cell layer.

SABCi-NS1.1 (115) (referred to as SA-BCi throughout) are a similarly immortalized cell line derived from normal human small airway epithelium basal cells. Like BCi cells, these cells retain differentiation capacity and when cultured at the ALI and form layers of ciliated, club, and secretory cells.

PBEC primary cultures of bronchial epithelial cells (PBEC) were isolated from human lung parenchymal samples obtained during cancer resection surgery. Patient samples were anonymised and harvested by pathology staff from the University of Southampton with Ethical approval granted by the Southampton and Southwest Hampshire Research Ethics Committee. Lung tissue samples were cut to 1mm pieces and cultured in collagen coated T75 flasks containing PneumaCult Ex plus media (comprised of PneumaCult Ex Plus Basal media, 1x PneumaCult Ex Plus supplement, 0.1% Hydrocortisone Stock Solution, 0.1% Gentamycin and 0.05% Amphotericin) to encourage PBEC outgrowth. In addition to this, the TL1831 sample was enriched with PBEC by taking brushes of visible large bronchioles present in the lung tissue sample. These cells were not used beyond 2 passages.

2.2.2 2D cell culture

Culture vessels were coated in a 1/50 dilution of collagen (PureCol) prepared in sterile water prior to propagation or maintenance of cells. The flasks/plates were placed in an incubator, 37°C, 5% Co2, for 30 minutes to allow the collagen solution to polymerise. Excess collagen was then removed and the flasks/plates were allowed to air dry under a tissue culture hood before use.

The basic steps of cell culture were the same for each cell type with differences in culture media only. VERO cells were maintained in Dulbecco's Modified Eagle media (DMEM) supplemented with 10% Foetal Calf Serum (FCS), 4mM L-Glutamine, 1% Penicillin/Streptomycin and 1% Sodium Pyruvate. BCi cells were maintained in PneumaCult Ex Complete culture media comprised of PneumaCult Ex basal media, 1x PneumaCult-Ex Supplement, 0.1% Hydrocortisone Stock Solution, 0.1% Gentamycin and 0.05% Amphotericin. SA-BCi cells as well as PBEC isolated from lung tissue were maintained in the same PneumaCult Ex Plus Complete media used for initial PBEC isolation.

In all cases, culture media was fully changed every 2 days until 60-80% confluent at which point the cells were passaged using a Trypsin solution to detach cells and then either split for continuous culture or seeded to new culture vessels as required.

2.2.3 Air liquid interface culture

Human airway epithelial cells from all three sources were differentiated using an Air liquid interface (ALI) culture method. Transwell membrane inserts (Corning) arranged in wells of a 24 well plate, were coated on the apical surface with a 1:10 dilution of collagen using the same method described in section 2.2.2. Prior to seeding, 800µl of PneumaCult Ex Plus culture media was added to the basolateral compartments and allowed to warm at 37°C, 5% CO^2. Cells were seeded onto the apical compartments at a density of $1x10^5$ - $2x10^5$ cells per insert in 200 µl of PneumaCult Ex Plus culture media. The media was replaced every 2 days and changed to PneumaCult ALI media in both compartments at the third media change. After 2 more media changes, the cells were “airlifted” by removing media from the apical compartment. Subsequent media changes replaced only the basolateral media. Differentiation required approximately 3-4 weeks after transfer to ALI media.

Beginning the second week after airlift, the apical compartments were gently washed in 200µl of HBSS to remove excess mucus. Differentiation was monitored both by visual observation of cilia movement by light microscopy and by using Trans Epithelial Electrical Resistance (TEER) to measure ionic permeability across the epithelial layer as an indicator of layer integrity. After initial experiments in which TEER was measured every 2 days post airlift, TEER was conducted once a week in conjunction with the weekly mucus washes. Briefly, 200µl of room temperature HBSS was added to the apical compartments of each insert and the pates were returned to the incubator for 10 minutes to allow equilibration and TEER was measured using a Millicell ERS-2 Voltohmmeter. TEER readings > 800 Ohms were considered suitable for further experimentation however in practice cells were not used below 1000 Ohms (330Ohms*cm^2).

2.3 Viral Infection

2.3.1 Viral Strains

The laboratory strain RSV-A2 was obtained from Virapur as pre-purified aliquots with an approximate concentration of $7x10^7$ PFU/Ml. Stock RSV-A2 was stored as 15µl aliquots at -80°C for ongoing use. All handling of RSV-A2 was performed at BSL-2.

2.3.2 2D culture infection

Submerged Vero, BCi or PBEC cultures were infected with RSV-A2 in 12 well culture plates (Corning, well surface area: 3.8 cm^2) or 4 well chamber slides (Nunc, Lab Tech, chamber surface area: 1.7 cm^2) both coated as described previously with a 1:50 dilution of collagen. Each 12 well plate was seeded at a density of 200,000 cells per well and chamber slides were seeded at 100,000 cells per chamber.

Vero cell monolayer infections were performed as 2-fold serial dilutions from an initial 1 in 10 dilution of RSV stock, shown in **Table 2**. In all 2D infections, viral dilutions were prepared in infection media (Vero: DMEM plus 4nM L-Glutamate, BCi/PBEC: PneumaCult Ex Basal media, 1% pen/strep) and mixed well before adding to the cells.

Table 2 RSV-A2 Serial Dilutions and per well concentration used for monolayer infections

	RSV (µl)	**Diluent (µl)**	**PFU (50µl)**
Dilution 1	50 Stock	450	$3.5x10^5$
Dilution 2	250 D1	250	$1.75x10^5$
Dilution 3	250 D2	250	$8.75x10^4$
Dilution 4	250 D3	250	$4.37x10^4$
Dilution 5	250 D4	250	$2.18x10^4$

RSV serial Dilutions and estimated MOI. Stock RSV-A2 virus was diluted 1 in 10 in cell specific infection media. This first dilution (D1) was then serially diluted 1 in 2 to produce 5 concentrations of RSV. In each 12 well plate experiments (plaque assay or RNA collection) 50µl of each dilution was added to 250µl of media per well or 25 µl of viral dilution per 125µl of media per chamber of a 4 well chamber slide.

The cells were washed twice with Infection media before adding 50 µl of each RSV-A2 dilution to 250 µl Infection media per well (12 well plate) or 25 µl RSV dilution to 125 µl media per well (Chamber slides). After pipetting to mix, the culture plates were placed in an incubator at 37°C, 5% Co2, for 2 hours, pipetting the media again 1 hour after infection. The Infection media was then discarded to Virkon and the wells were washed twice in infection media. Finally, 1ml (per well) or 500 µl (per chamber) of Post infection media (Vero Infection medium + 5% FCS, BCi Infection media + 0.02% BSA) was added for a final incubation period of 72 hours.

2.3.3 RSV Plaque Assay

The modified plaque assay described by McKimm-Breschkin JL (116) was used with minor modification. Briefly, Vero or BCi cells were grown to near confluence in 12 well culture pates and infected with 2-fold serial dilutions of RSV as described in section 2.3.2. After the 2 hour incubation period, infection media was replaced with 1.5 ml of overlay media comprised of either agarose (0.3% or 0.5%) prepared in VERO Post infection medium or an overlay media of microcrystalline cellulose Avicel™. The plates were then returned for the incubator and left undisturbed for 6 days before fixing the cells by adding 2ml of 4% PFA through the Overlay and incubating at room temperature for 30 minutes. The overlay and fixative were then removed, and the cells were stained with either 0.05% Neutral Red or Crystal violet. Plaques are observed as discrete unstained regions of infected, metabolically inactive or dead cells which do not take up the dye.

2.3.4 ALI culture infection

ALI cultures were infected with RSV using the same basic protocol for both BCi and primary cell models. Briefly, the apical surface of each well was washed using 200µl of HBSS and the basolateral medium was replaced with 800µl of ALI Infection media (PneumaCult ALI basal media, 1X

Pneumacult ALI supplement, 1% pen/strep, 0.02% BSA). RSV stocks were prepared in HBSS to the specified concentrations and added to the cells in 50µl volumes. The bulk of experiments used 2.5 µl of RSV per well, approximately $1.75x10^5$ PFU or an MOI of 1 for an estimated cell density of $1.8x10^5$ cells. Uninfected, HBSS only, controls were included with each infection. Viral concentrations for all volumes of stock used in ALI infections are shown per well in **Table 3.**

Table 3 RSV-A2 concentrations used for ALI infection

RSV (µl)	Diluent (µl)	PFU (50µl)
20	30	$1.40x10^{06}$
10	40	$7.00x10^{05}$
5	45	$3.50x10^{05}$
2.5	47.5	$1.75x10^{05}$
1.25	48.75	$8.75x10^{04}$
0.625	49.375	$4.38x10^{04}$

RSV-A2 was used at fixed volumes for each infection. In practice, the above concentrations were achieved using serial dilution for volumes needed for each experiment. Results for ALI infections are described for each per well volume of RSV-A2. Apical RSV-A2 infections were performed in 50µl volumes comprised of the volume of RSV and HBSS diluent described here. PFU = Plaque forming units

The plates were returned to the incubator at 37°c and 5% Co2, with gentle pipette mixing of the apical viral solution after 1 hour. After 2 hours, the viral solution was discarded to Virkon and the apical surfaces were washed with 200 µl of HBSS twice, pipetting gently before removal of each wash. Samples were collected at 2hpi as baseline measures of RSV infection. The remaining cells were returned to the incubator for a further 70 hours for a total 72 hour infection.

2.4 ICC/IHC and Immunofluorescence staining

Samples were prepared for both chromogenic and fluorescent imaging of RSV or SARS-CoV2. Monolayer cell cultures were imaged in 4 well chamber slides (NunTec) which allows culture, infection, fixation and staining to occur directly on the slide used for imaging. The cells were fixed in 4% Paraformaldehyde (PFA) for 30 minutes before washing gently in PBS. The cells were then permeabilised with 0.1% triton X-100 in PBS for 15 minutes at room temperature followed by

incubation for 1 hour in peptide blocking buffer (1% BSA in PBS plus 0.1% tween20). Primary antibodies (anti-RSV-A2 G – Rabbit Pab GTX70381) were prepared in the same blocking buffer and added at 1:500 dilution overnight at 4°C. After washing the cells (3x 5min) in PBS-T (0.1% tween), fluorescent labelled secondary antibody (Goat anti-Rabbit AF-568) was added in blocking buffer (1:500 dilution) for 1 hour at room temperature. After washing (3x 5min) unbound antibody with PBS-T, the nuclear dye DAPI was added for 5 minutes (1:1000 in PBS). After a final wash in PBS, the chambers were removed and a coverslip mounted using Mowiol. The slides were imaged once fully dry using an AxioskopII fluorescent microscope.

ALI samples were fixed by adding 1ml of 4% PFA gently into the apical compartment of HBSS rinsed inserts, allowing the solution to overflow into the basal compartment. After 30 minutes at room temperature, the PFA was removed and the inserts were placed into 70% ETOH for transfer to the histochemical research unit (HRU) for paraffin embedding. Wax embedded transwell membranes were cut to 5µM sections using a microtome and allowed to dry overnight on 3-Aminopropyltriethoxysilane (APES) coated microscope slides. Wax was then removed in XTF Clearing Agent (CellPath) and the samples were rehydrated through grades of alcohol. Antigen retrieval was conducted by heating samples for 25 minutes in a microwave at 50% power in Citrate buffer (10mM Citric acid in dH2O. pH 6). For fluorescent imaging, the same permeabilization, blocking and antibody incubation steps described for monolayer slides were then used for these samples. SARS-CoV2 staining was conducted using an in house rabbit polyclonal anti-SARS-CoV2-N antibody at 1:400 dilution.

For chromogenic staining with DAB, dewaxed and rehydrated sections were treated with 0.5% Hydrogen peroxide in Methanol for 10 minutes to block any endogenous peroxidases. This also removes the need for further permeabilization. The same Antigen Retrieval, peptide blocking, and primary antibody incubation steps were then followed. HRP-conjugated anti-rabbit secondary antibody (1:1000) was added for 1 hour at room temperature. After washing the slides in PBS, the

chromogenic substrate was added for 5 minutes. The slides were then rinsed in PBS and placed in running tap water for 5 minutes before briefly counterstaining with Haematoxylin and rinsing in running tap water for a final 5 minutes. The slides were then dehydrated through increasing grades of alcohol followed by XTF clearing reagent and coverslip mounting.

2.5 qPCR

2.5.1 Sample Collection and RNA Isolation

To collect samples for RNA isolation, transwell or monolayer cultures were rinsed to remove unbound RSV or media before adding 1ml of Qiazol reagent per well or transwell insert. In ALI experiments, the Qiazol was allowed to overflow from the apical compartments ensuring both sides of the membrane were in contact with the reagent. After a few minutes the Qiazol/cell solution was transferred to sterile RNAse free 1.5ml Eppendorf tubes, pipetting each sample carefully during collection. Samples were left at room temperature for 5 minutes to ensure full dissociation before being stored at -80c or proceeding to RNA isolation.

RNA was isolated from Qiazol samples using phenol chloroform extraction. Briefly, each 1ml Qiazol sample was vigorously mixed with 0.2ml of Chloroform for roughly 15 seconds. After resting for 5 minutes at room temperature, the samples were centrifuged at 12,000xg for 15 minutes at 4°C to induce phase separation. The top aqueous phase was transferred to pre-prepared 1.5ml Eppendorf's containing 0.5ml Isopropanol and 20ug of Glycogen. The samples were briefly vortex mixed then placed at -20°C to incubate for 20 minutes before repeating centrifugation to pellet RNA. The supernatant was removed and the visible RNA pellet was washed twice with 1ml of 75% Ethanol, centrifuging at 12,000xg for 10 minutes at 4°C in between washes. After removing the final wash, residual Ethanol was allowed to evaporate before dissolving the RNA pellet in 15µl of Nuclease Free

water. RNA samples were then quantified using NanoDrop 1000 software (Thermo Fisher Scientific) and stored at -80°C.

2.5.2 Reverse Transcription

RNA was reverse transcribed to complementary DNA (cDNA) using reagents from Applied biosystems. For each sample, 250ng of RNA was prepared to a volume of 13.2µl in RNAse-Free water. This was added to 6.8 µl of Reverse transcription master mix containing: 2µl 10 x RT random primers, 2µl 10 x RT buffer, 0.8µl deoxyribonucleotide triphosphate (dNTP) mix, 1µl RNAse inhibitor and 1µl Multiscribe Reverse Transcriptase. The final 20µl volumes were then exposed to thermal cycling at 25°C for 10 min, 37°C for 2 h and 85°C for 5 min using a Tetrad DNA Engine Thermal Cycler (MJ Research/Bio-Rad Laboratories CA, USA). Samples of cDNA were then diluted 1:10 in RNAse-Free water and stored at 4°C.

2.5.3 Quantitative PCR (qPCR)

Quantitative PCR (qPCR) was performed using TaqMan® probe–based assays (**Table 4**) in 384 well pates. A 1 µl aliquot of each diluted cDNA sample was added to 4 µl of primer master mix (2.5µl of TaqMan Universal Master Mix II (Applied Biosystems), 1.25µl of RNAse-free water (Sigma) and 0.25µl of appropriate TaqMan primers) per well. Reactions were performed in duplicate and amplified using the following cycling parameters: 95°C for 10 min and 40 cycles of 95°C for 15s and 60°C for 1 min using a 7900HT Fast Real-Time PCR System.

Table 4. qPCR TaqMan probe assays

Primer	Fluorescent label	TaqMan Assay ID
HPRT1	VIC	Hs502800695_m1
GAPDH	VIC	Hs02758991_g1
RSV-N	FAM	1305382 A7
CXCL10	FAM	Hs00171042_m1

TaqMan assays used in individual qPCR single-plex experiments were purchased from ThermoFisher. All primers were used at the recommended volumes with TaqMan Universal Master Mix II (Applied Biosystems).

Relative gene expression was determined using the comparative Ct or ΔΔCt method proposed by Livak et al (117). By this method, averages of technical replicates of all experimental and baseline Ct values are first normalised to an endogenous control or housekeeping gene (HKG) before the resulting baseline ΔCt values are subtracted from the experimental ΔCt as shown below:

$$\Delta\Delta Ct = (Experimental\ GOI - Experimental\ HKG) - (Baseline\ GOI - Baseline\ HKG)$$

In this work relative expression was calculated using either uninfected or 2hpi samples as baseline measurements. As described in the livak method, final fold changes are calculated as: $2^{-(\Delta\Delta Ct)}$

2.5.4 Statistical Analysis

Statistical analysis for in vitro experiments was performed using GraphPad Prism (version 9.2 GraphPad Software, San Diego, USA) using non-parametric Wilcoxon matched-pairs signed rank test to compare fold changes between infected and uninfected samples or between RSV infection concentrations. Statistical significance was determined as $p<0.05$ however no comparisons met this threshold.

All Bioinformatics statistical analysis was performed using the packages described in the Bioinformatics methods.

Chapter 3 Comparison of viral gene expression datasets

3.1 Introduction

Databases containing published transcriptomics data such as the NCBI GEO repository are an invaluable resource for researchers looking to explore new hypotheses on existing data. Additional data sets are continually added with an increasing expectation that raw data will be made available to the research community with new publications. Whilst independent control of experimental parameters is lost, this collaborative approach to research and data sharing reduces the need for experiments to repeated unnecessarily and allows for data generated from multiple sources to be combined and compared in a single new analysis.

In this project, the NCBI GEO repository was used to obtain transcriptomic data generated from airway epithelial cell models of common respiratory viral infections. To allow for differential expression analysis, time-matched uninfected controls needed to be included with each data set as well as sufficient replicate samples of each experimental condition. The search was further restricted to only human ALI models due to the improved biological relevance of this approach over 2D cultures. While this greatly decreased the number of available data sets, maintaining consistency in experimental design was prioritized to minimise the impact of variables other than infection on the analysis.

The two data sets ultimately selected for this project describe microarray gene expression changes in ALI cultures infected with four respiratory viruses, two per data set. They were ideal for comparative analysis as both data sets were generated from experiments using primary human bronchial epithelial cells in ALI culture with similar infection protocols. Both data sets also used the Agilent single colour Human Genome Microarray - G4112F platform, which removed an additional source of technical variation between datasets. An overview of the experimental models used by the

publishing authors as well as citations for any associated published work are included prior to the results subsection for each data set.

Data from each viral infection were processed as independent experiments followed by comparison of gene expression differences within data sets and across viral infections. Direct comparison of gene expression at each time point in response to different viruses was not performed due to the expected impact of experimental variables on the results. Importantly for this project, antiviral and inflammatory gene expression in infected samples would be impacted by differences in viral load. This could vary throughout an infection due to factors including the initial viral inoculum and the health or phenotype of the target cells, in addition to incubation time.

Differentially expressed genes from each experiment were used for pathway analysis and these results were contrasted across viruses, using time points with the greatest magnitude of change where relevant. The data were additionally filtered to explore any patterns in gene expression with roles in T cell activation and recruitment pathways. This following chapter describes the steps taken to fulfil the first and second aims of this project; analysis and comparison of transcriptomic data generated from virally infected PBEC ALI models of respiratory viral infection.

3.2 Results

3.2.1 GSE32138 Analysis

GSE32138 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE32138) produced by Ioannidis et al, explores infection with two common respiratory viruses, RSV and a seasonal strain of IAV (H3N2). For these experiments, the authors utilized Human airway tracheobronchial epithelial cells, obtained from airway specimens taken during resected at lung transplantation. Isolated epithelial cells were transferred to Transwell inserts for ALI differentiation over 4-6 weeks.

The differentiated cells were infected for 2h with either Influenza- A/Udorn/72 H3N2 viruses at MOI-1, relative to the apical cells, or the laboratory strain RSV-A2 at MOI-5. The H3N2 infection was incubated for 24 hours post infection while RSV-A2 was incubated for 48 hours. Mock infected controls were simultaneously exposed to the same experimental conditions and incubation times without virus added.

Cells from two separate human donors were used to create four samples per condition, resulting in sixteen arrays in total. Infected cultures used two samples from each donor and mock-infected cultures were all from the same donor. Information on which samples were generated from each donor was not provided with the raw data files, however the authors state that no indication of donor effects was observed (*Published data from this analysis* (105)).

3.2.1.1 Pre-Processing and Quality control Testing

The raw array intensity files from this data set were appropriately normalised using the *NEQC* background correction and quantile normalisation. Boxplots of the log transformed raw expression data show minor variations in probe intensity between arrays within each group and larger differences across groups prior to normalisation. Following background correction and quantile normalisation the spread of the data is even across all arrays **Figure 3-1**.

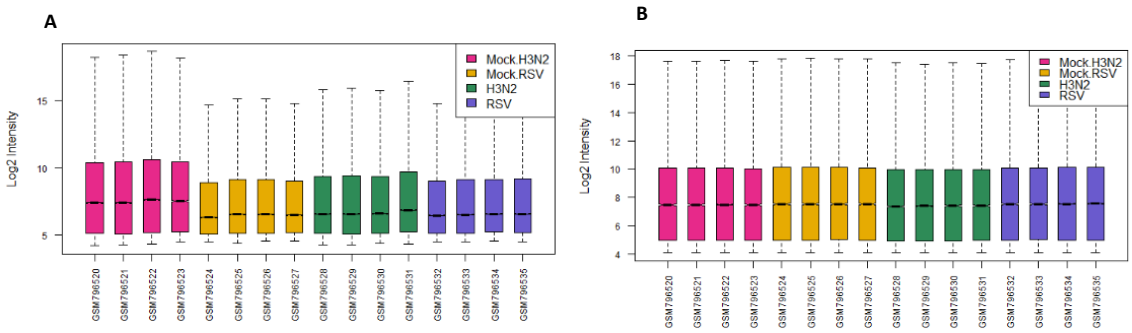


Figure 3-1 GSE32138 array intensity values before and after NEQC normalisation

GSE32138 array intensity values were normalised using the NEQC function from the limma package. This approach uses the control probe data included with the raw files to perform background correction followed by quantile normalisation. Intensity values for all arrays are visualised as boxplots both before (A) and after (B) normalisation. Arrays are coloured according to group membership.

The structure of the normalised data was visualised using multidimensional scaling (MDS) and hierarchical clustering **Figure 3-2**. The arrays clustered into 4 distinct groups based on experimental treatment. Using MDS, arrays in both infection groups were separated from arrays in the mock-infected treatment groups along the x-axis, or first dimension which explains 60% of the data variation. H3N2 infection arrays (Green) are particularly distinct along this axis while RSV infection arrays had greater variation from controls along the y-axis or second dimension which explains 18% of variation. The 4 arrays belonging to each mock-infection group, Mock.RSV (Yellow) and Mock.H3N2 (Pink), also clustered together based on treatment group with slight separation along dimension 2.

Hierarchical clustering generated a similar pattern; individual arrays clustered by group membership and the H3N2 infected samples were separated from the other arrays at the first branch of the dendrogram. RSV infected samples were clustered with the least distance from mock infected samples, showing the most similarity to arrays from the Mock.RSV group. No outliers were present in the data.

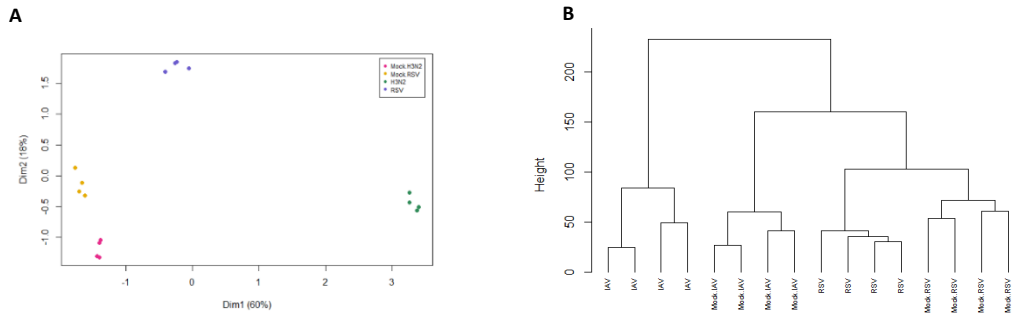


Figure 3-2 GSE32138 arrays cluster by treatment group

A) MDS plot showing the variation between normalised intensity values from the GSE32138 data. Spots represent individual arrays coloured by viral agent (RSV = Blue, IAV = Green) used during PBEC infection or time matched mock infection (mock.RSV = Yellow, Mock IAV = Prink). The distance between points approximates the greatest log2FC between samples, B: Dendogram produced from hierarchical clustering of the same data using ward linkage.

3.2.1.2 H3N2 and RSV: Differential expression analysis

Differential expression (DE) analysis of infected arrays relative to mock infected controls identified a total of 4540 genes affected by H3N2 and only 674 genes affected by RSV beyond the chosen threshold for significance (Adjusted P value 0.05) and effect size (log2FC > 1 or >-1). As displayed in **Figure 3-3**, 520 DE genes were common to each infection, 4020 genes were DE by H3N2 alone and just 154 genes were uniquely DE by RSV.

H3N2 infection resulted in close to equal numbers of upregulated (2038) and downregulated (2457) genes while RSV infection resulted in 530 upregulated genes and only 144 downregulated genes.

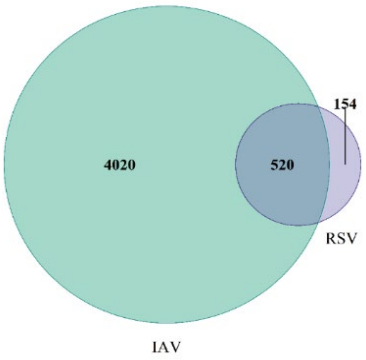


Figure 3-3 H3N2 resulted in differential expression of 4020 genes not impacted by RSV.

Venn diagram showing the number of combined upregulated and downregulated genes from each viral infection. Generated from time matched DE analysis.

The overall magnitude of DE was also larger in response to H3N2 infection than it was for RSV. Volcano plots showing the log2FC in both directions generated for time matched DE analysis of each viral infection are shown in **Figure 3-4**. Each point represents a single gene positioned on the plot based on log2FC (x-axis) and significance p. values (y-Axis). Vertical and horizontal bars respectively show thresholds of effect size and significance. The few genes downregulated by RSV have relatively low log2FC values as well as significance. In contrast, the H3N2 data shows a broader spread both values and a concentration of genes with higher log2FC.

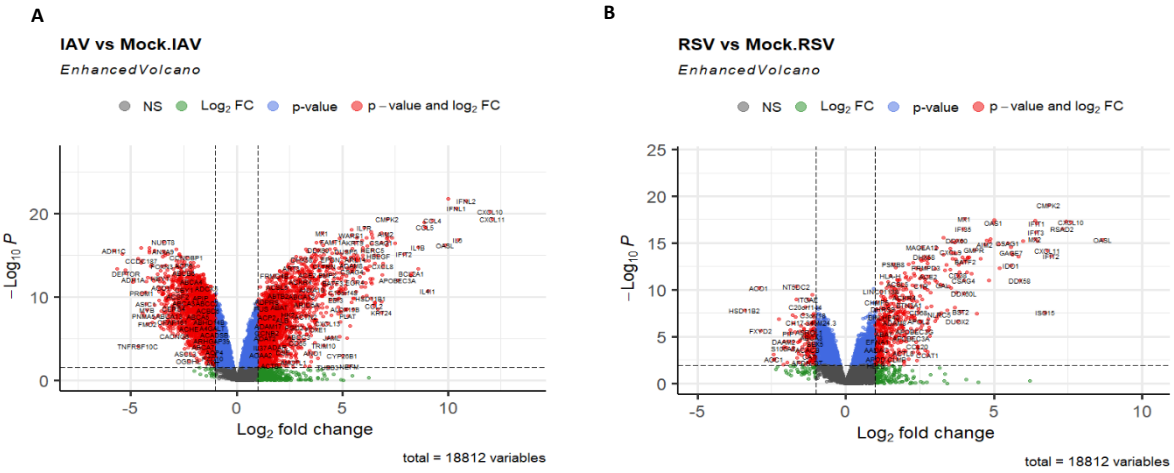


Figure 3-4 GSE32138 Volcano Plots: H3N2 has a larger and more varied impact on gene expression than RSV

GSE32138 expression data were normalised using NEQC methods and DE analysis was performed using LIMMA for infected samples relative to mock infected controls. The result of this analyses for IAV (H3N2) A. and RSV (B) are visualised as volcano plots with significance P values represented in the -Log10 scale to show genes with greater significance as higher values on the Y axis. Vertical dotted lines indicate log2FC threshold of 1and the horizontal line marks the equivalent P value for an adjusted P value cut of 0.05.

Significantly DE genes identified from each viral infection were taken for pathway annotation and ORA using the clusterProfiler package. The background (Universe) gene set was comprised of all genes with FDR < 0.05 without filtering by effect size. From 2038 genes upregulated by H3N2 and 530 upregulated by RSV, 1854 and 477 were mapped to GO:BP terms respectively. The most significantly enriched terms were associated with the immune response for both viral infections, as shown in **Figure 3-5 (A and B).**

H3N2 infection caused more significant enrichment of cytokine focused GO:BP terms such as "cytokine-mediated signalling pathway" (GO: 0019221) than RSV. There were 152 genes (FDR $8.67x10^{-31}$) from this pathway term upregulated by H3N2 and only 58 (FDR $2.33x10^{-13}$) upregulated by RSV. Specific antiviral terms were enriched with very similar (or slightly stronger for RSV) FDR adjusted P values by both viruses despite H3N2 resulting in upregulation of more genes from this term, for example H3N2 upregulated 106 genes (FDR $7.69x10^{-24}$) from "defence response to virus" (GO:0051607) while 73 genes (FDR $2.18x10^{-29}$) were upregulated by RSV. This equates to roughly 50% of all mapped genes upregulated by RSV.

Genes downregulated by H3N2 were significantly enriched for pathway terms relating to cilia, for example "cilium organisation" (GO:0044782) and "microtubule-based movement" (GO:0007018). ORA of the genes downregulated by RSV did not identify and significantly enriched terms (FDR < 0.05), however genes inhibited by RSV were mapped to terms relating to cellular metabolism such as "lipid metabolic process" (GO:0006629) and "organic acid metabolic process" (GO:0006082).

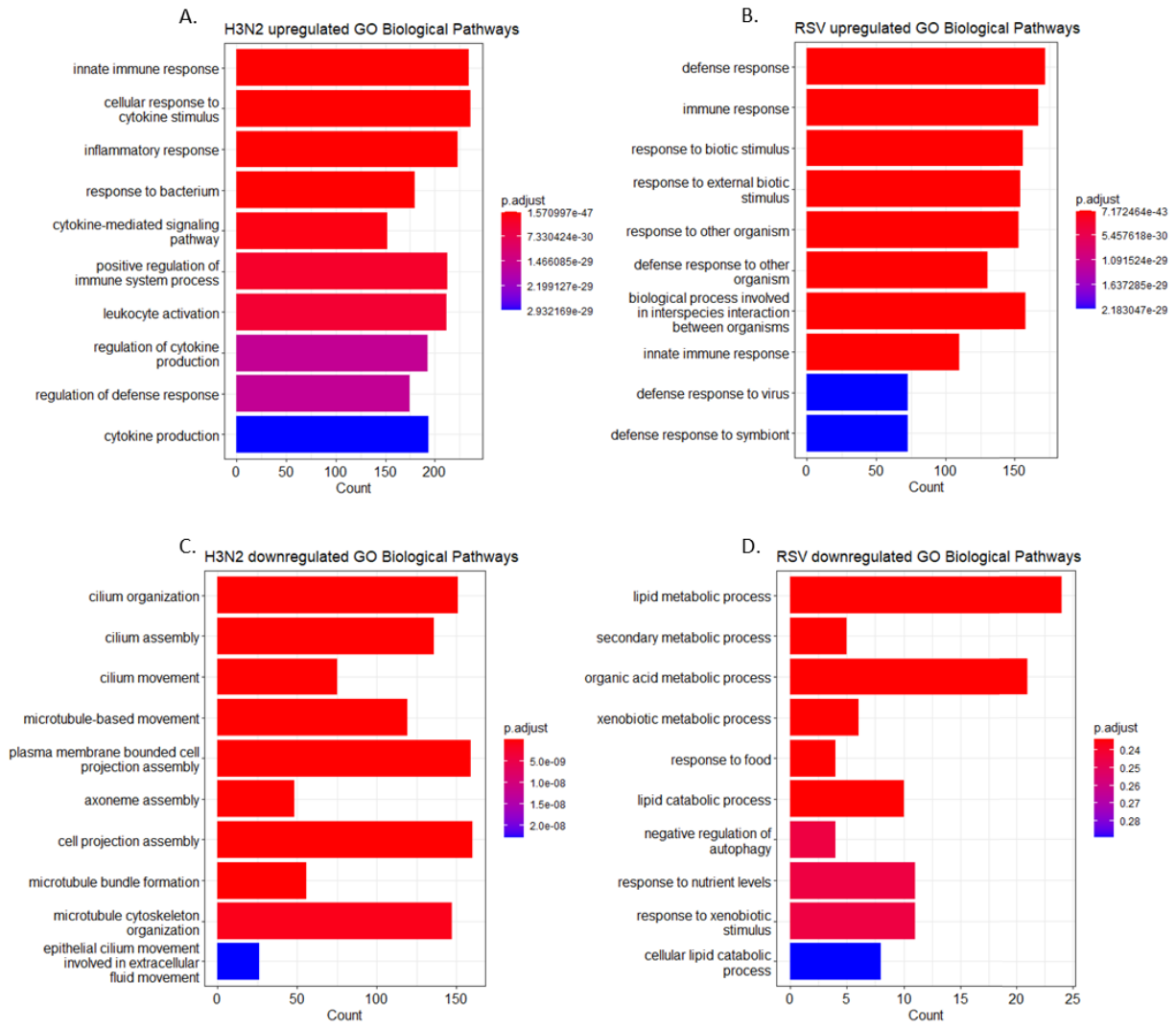


Figure 3-5 Over representation analysis using clusterProfiler identified upregulation of immune system pathways and downregulation of cilia formation and cellular metabolism by H3N2 and RSV

GSE32138 expression data were normalised using NEQC methods and analysed for DE using Limma. Significantly up- and downregulated genes from each viral infection were annotated using clusterProfiler with significant genes, with no effect size filter, used as the background gene set. Plots are coloured by FDR adjusted P value. A & B show the 10 most significantly enriched GO:BP terms from genes upregulated by H3N2 and RSV respectively. C & D show GO:BP terms enriched for downregulated genes.

Genes relating to antiviral immunity were among the most highly upregulated by both viruses with many IFN related genes showing comparable levels of DE by RSV and H3N2. H3N2 additionally resulted in significant upregulation of cytokine encoding genes, which were not DE by RSV. Normalised expression values for all 82 genes, from either DE gene list, mapped to "defence response to virus" were visualised as a heatmap to compare data from RSV and H3N2 infection

Figure 3-6. Data from both infected and mock-infected arrays for each experimental condition are

shown with rows (genes) scaled as Z-Scores, showing variation around the mean expression value of each gene. Hierarchical clustering of rows as well as columns (arrays) was performed using Wards algorithm. Replicate arrays for each experimental treatment group clustered together, including mock-infected arrays.

Log2FC and adjusted P values from time matched differential expression analysis for a selection of these genes are included in **Table 5.** There was comparable upregulation of most type-I IFN inducible genes in response to infection with both viruses including the interferon regulatory factor genes (*IRF7* and *IRF9*), interferon stimulated genes (*ISG15* and *ISG20*) and OAS and IFIT family members. There was a larger difference in expression of both type I and type III IFN genes, specifically, RSV had no impact on any type I IFNs while H3N2 caused DE of several variants of IFNa (*IFNA4, IFNA21, IFNA8, IFNA14, IFNA5, IFNA10*) as well as *IFNE* and *IFNB1.* The log2FC for these genes was lower than for the type III IFNs however as H3N2 had a particularly strong impact on expression of *IFNL1*, *IFNL2* and *IFNL3*. All three of these genes were also DE by RSV, to a lesser extent.

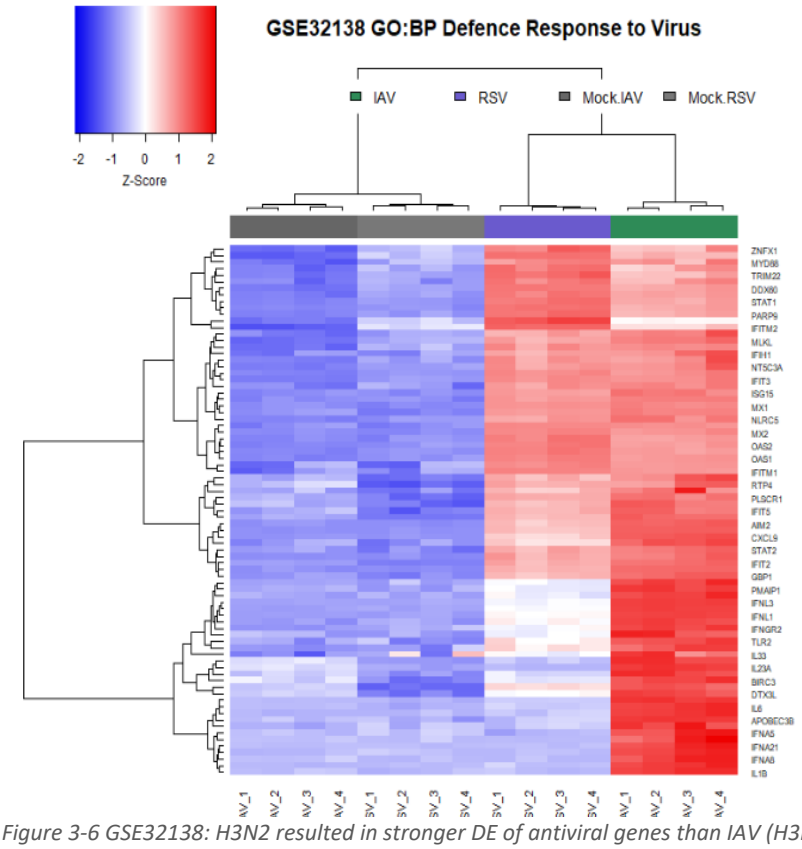


Figure 3-6 GSE32138: H3N2 resulted in stronger DE of antiviral genes than IAV (H3N2)

GSE32138 expression data were normalised using NEQC methods and analysed for DE using Limma. Significant DE genes with relevance to "Defence Response to Virus" were identified using ClusterProfiler. This heatmap shows normalised expression data for all arrays from this data set. Both rows and columns were clustered using ward linkage. Group membership of each of 4 replicate arrays are shown as coloured bands for each infection group.

Table 5 GSE32138 Time matched DE of GO:BP Defence response to virus

	RSV		H3N2	
SYMBOL	Log2FC	FDR	Log2FC	FDR
OASL	8.68	4.86E-13	9.87	3.10E-14
IFI44L	7.44	1.62E-14	6.18	3.65E-14
RSAD2	7.31	5.28E-14	8.94	8.27E-16
IFIT2	6.93	1.13E-11	7.93	1.75E-13
ISG15	6.72	1.94E-06	6.4	8.87E-11
OAS2	6.47	5.64E-13	6.12	6.36E-13
IFIT1	6.44	1.76E-14	6.41	7.08E-15
IFIT3	6.42	1.03E-13	7.35	7.08E-15
OAS3	6.39	1.62E-14	5.92	1.13E-14
STAT1	4.86	6.10E-13	4	3.96E-13
IRF7	4.16	1.47E-11	4.25	4.72E-12
IFITM1	3.73	2.55E-08	3.69	3.08E-09
IFNL1	3.71	5.96E-12	10.39	2.05E-17
IFNL2	3.3	8.51E-12	10.87	4.13E-18

	RSV		H3N2	
SYMBOL	Log2FC	FDR	Log2FC	FDR
ISG20	3.29	3.13E-13	3.78	1.71E-14
IFNL3	2.8	1.44E-11	10	4.13E-18
IRF9	2.62	2.25E-11	1.68	8.34E-10
STAT2	2.45	8.94E-07	2.44	1.39E-09
TLR3	2.14	1.06E-07	2.11	1.27E-08
TLR2	1.31	8.62E-05	2.88	1.31E-09
IFNE	0.78	3.56E-04	3.09	8.76E-12
IFNGR2	0.67	5.39E-05	2.75	5.39E-13
TNFAIP3	0.64	1.10E-02	5.67	7.08E-15
IFNB1	0.51	4.99E-02	7.49	1.82E-14
TRIM15	0.17	5.83E-01	5.57	2.24E-12
IFNA8	0.01	9.77E-01	3.22	2.22E-11
IFNA14	-0.07	8.76E-01	3.14	4.88E-09
IFNA4	-0.13	6.44E-01	5.49	3.35E-13

Time matched DE analysis was performed on NEQC normalised expression data, the result of this analysis for genes with relevance to "Defence Response to Virus" in either gene list, identified using ClusterProfiler. are shown as log2 fold change (log2FC) with significance determined by FDR adjusted P value. AdjP values are shaded by significance with a 3-point scale. Yellow = >0.05.

3.2.1.3 H3N2 and RSV: PBEC – T cell gene expression

Enrichment of T cell related genes were identified by filtering for relevant GO:BP terms from the total ORA results for upregulated DE genes from each viral infection **(Figure 3-7)** . The significance values assigned to these terms were lower than for the more general terms discussed previously, however, there was moderately significant representation of terms including “T cell migration” (GO:0072678) and “T Cell activation “(GO:0042110) in the H3N2 generated gene set, the latter being mapped to 121 genes significantly DE by this virus. RSV infection resulted in weaker enrichment of most of these terms than H3N2, especially concerning T cell migration” which was not significantly overrepresented in the RSV gene list with only 7 DE genes mapped to this term. Both viral infections resulted in similar, weak, enrichment of GO:BP “antigen processing and presentation” (GO:0019882) and “negative regulation of T cell activation” (GO:0050868), despite RSV resulting in considerably fewer DE genes overall. **Table 6** shows FDR values calculated from ORA of these T cell terms.

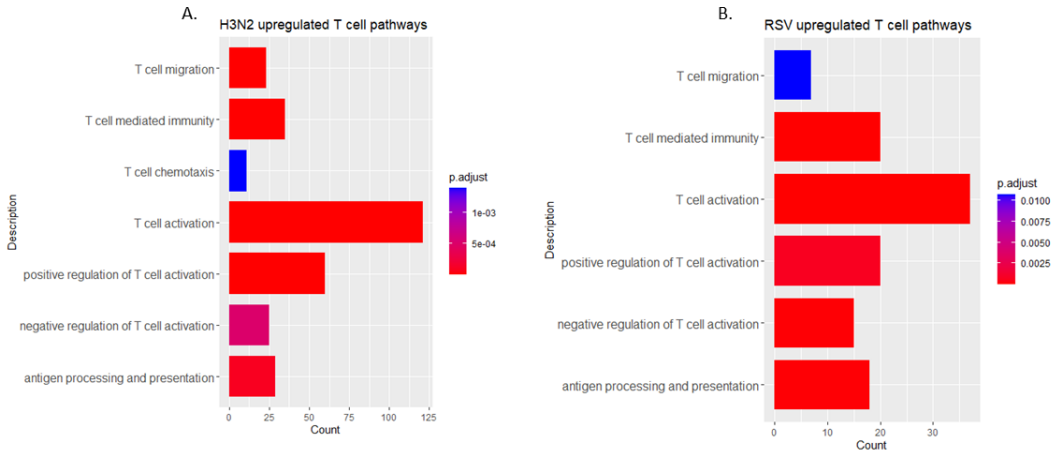


Figure 3-7 GSE32138 T cell recruitment and activation pathways were more significantly overrepresented in the H3N2 upregulated genes

Differentially expressed genes identified from the GSE32138 data set using limma, were annotated using over representation analysis with the clusterProfiler package. The normalised expression data were analysed for DE genes using limma. The 2038 genes significantly upregulated by H3N2, and 530 genes significantly upregulated by RSV, were analysed relative to all genes with adjusted p values < 0.05 in each infection analysis. The resulting GO biological process pathway terms identified for each virus were filtered to show pre-determined T cell activation or recruitment terms if they were enriched for each gene list. H3N2 infection (A) produced significant results for all terms. RSV infection (B) was less significant.

Table 6 GSE32138 H3N2 and RSV DE genes were enriched for T cell activation and recruitment GOBP terms

		H3N2 (1854 genes)		RSV (477 genes)	
ID	**Description**	**Count**	**FDR**	**Count**	**FDR**
GO:0042110	T cell activation	121	2.71E-18	37	8.91E-06
GO:0002456	T cell mediated immunity	35	2.08E-09	20	1.26E-06
GO:0050870	positive regulation of T cell activation	60	2.56E-09	20	7.22E-04
GO:0072678	T cell migration	23	3.06E-07	7	1.07E-02
GO:0019882	antigen processing and presentation	29	9.34E-05	18	8.27E-05
GO:0050868	negative regulation of T cell activation	25	5.13E-04	15	9.16E-05

T cell recruitment or activation related Gene ontology biological process terms were identified from ORA performed for genes upregulated by H3N2 and RSV using the clusterProfiler R package. The number of genes mapped to each term are shown along with the enrichment significance p value, adjusted for multiple hypothesis testing (FDR)

The 23 genes matched to "T cell migration" in the H3N2 data were reduced to 6 with relevance to PBEC – T cell recruitment. There were no additional unique genes relating to this subject in the RSV ORA. Expression data for these 6 genes are shown as a heatmap in **Figure 3-8** and the corresponding differential expression statistics are shown in **Table 7.** Both viruses had a particularly strong impact on the chemokine genes *CXCL10* and *CXCL11*, which were among the most highly DE in each analysis. However, the log2FC for all genes in this category were higher for H3N2 than they were for RSV. *CCL5* was the most highly upregulated CC chemokine in the RSV analysis followed by CCL20. Log2FC for these genes were more than twice as high in response to H3N2. Finally, RSV had no significant impact on *CCL3* or *CCL2*, both of which were upregulated by H3N2, *CCL3* with particular abundance.

The 114 genes matched to "T cell activation" from one or both gene lists were reduced to 28 with relevance to this project. This PBEC – T Cell Activation gene set was further categorised into two groups based on whether the gene products are secreted cytokines or surface bound e.g., co-stimulatory molecules or MHC receptors. There was comparable low upregulation of most MHC encoding genes with the exception of *HLA-DPA1*, which was not upregulated by H3N2. A heatmap generated from expression data for these genes (**Figure 3-9**) shows that this might not be a reliable finding as the mock infected H3N2 baseline samples show mixed intensity values for this gene.

The remaining genes in this category were DE by H3N2 to a greater extent than they were by RSV. Only 5/14 cytokine genes were significantly DE by RSV, and those 5 were upregulated with relatively

small log2FC compared to H3N2. This difference was most substantial for *IL6*, *IL1B* and *TNF*, all of which were not altered beyond the significance threshold by RSV but were highly upregulated by H3N2 **(Table 8).**

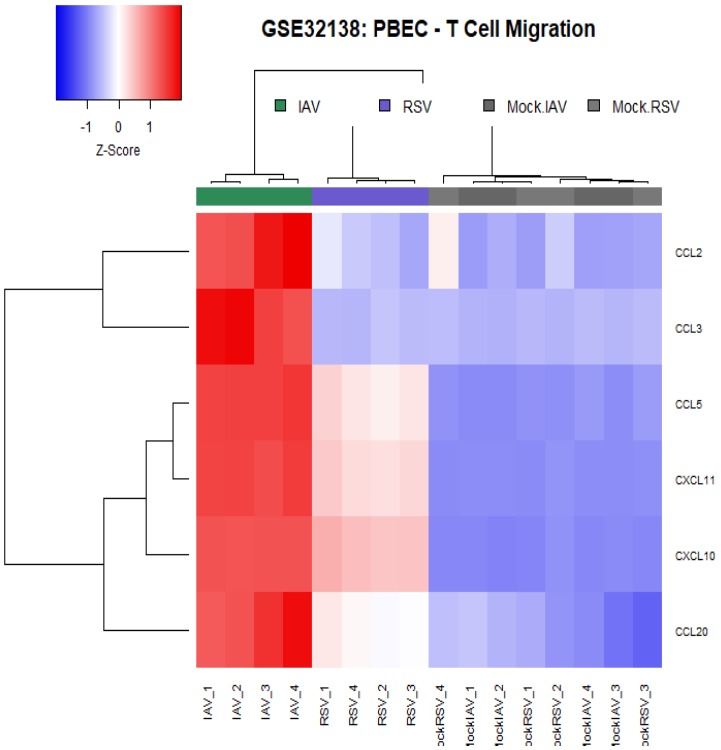


Figure 3-8 H3N2 resulted in consistent upregulation of genes in the PBEC - T cell Migration gene set

Heatmap generated from normalised GSE32138 array intensity values for genes classified into the PBEC – T cell Migration gene set. Rows (genes) have been scaled as Z-Scores showing variation of expression around the mean expression for each gene, Blue shaded boxes indicate genes with up to 2 standard deviations lower than the mean, red boxes indicate genes with expression up to 2 standard deviations above the mean. All 4 replicate arrays are included and are labelled at the base of the plot. Arrays as well as genes were clustered using Wards linkage. Coloured bars at the top of the plot correspond to each treatment group and show good clustering of arrays based on treatment.

Table 7 Pairwise differential expression of GSE32138 RSV and H3N2 infected PBEC relative to time matched controls for gene in the PBEC - T cell Migration gene set

	RSV		H3N2	
SYMBOL	**log2FC**	**FDR**	**log2FC**	**FDR**
CXCL11	6.8	3.44E-12	12.09	1.95E-16
CXCL10	7.6	1.62E-14	11.97	3.35E-17
CCL5	4.01	2.50E-11	8.9	8.13E-16
CCL3	0.05	9.39E-01	8.23	1.47E-11
CCL20	2.43	1.44E-03	6.53	5.55E-09
CCL2	-0.06	9.52E-01	6.49	1.14E-08

Log2 fold change (log2FC) values were calculated using pairwise differential expression analysis of RSV or H3N2 infected PBEC ALI relative to mock infected controls. Results are shown for genes mapped to the GO:BP term "T cell migration", with values arranged by decreasing log2FC from the H3N2 data. Adjusted P values (FDR) are corrected for false discovery of multiple hypothesis testing using eBayes (limma). Values are coloured by significance in a 3-point scale with yellow indicating AdjP > 0.05.

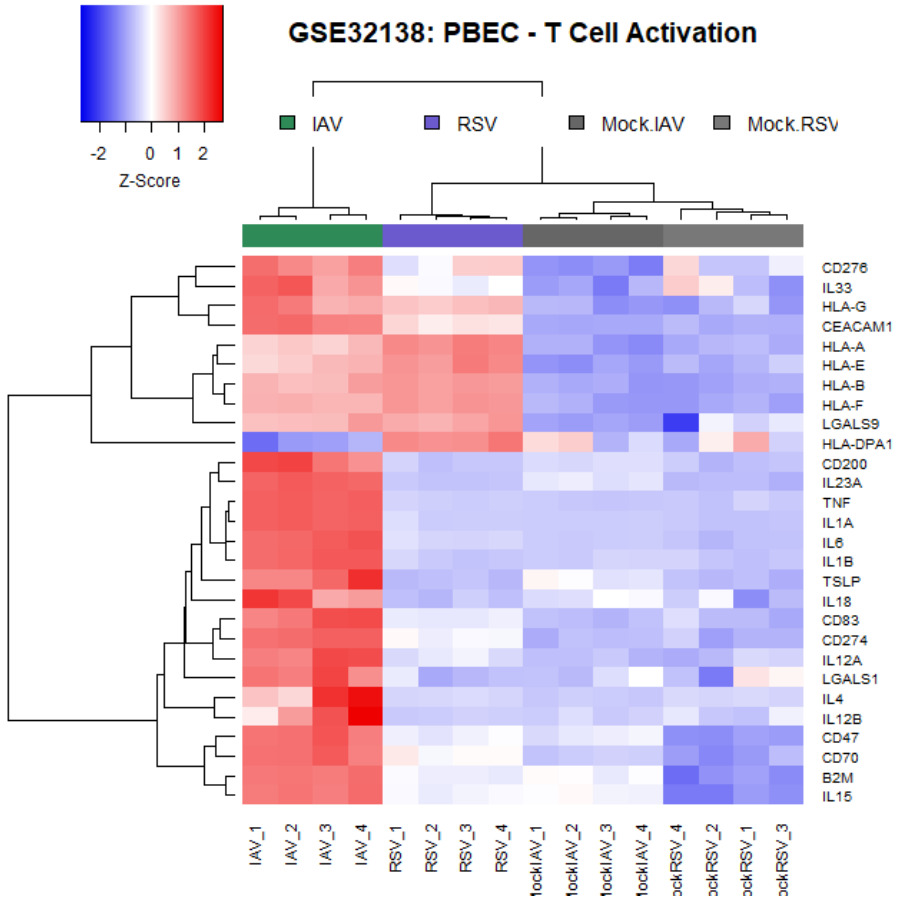


Figure 3-9 RSV and H3N1 expression of PBEC - T cell activation gene set

Heatmap generated from normalised GSE32138 array intensity values for genes classified into the PBEC – T cell Activation gene set. Rows (genes) have been scaled as Z-Scores showing variation of expression around the mean expression for each gene, Blue shaded boxes indicate genes with up to 2 standard deviations lower than the mean, Red boxes indicate genes with expression up to 2 standard deviations above the mean. All 4 replicate arrays are included and are labelled at the base of the plot. Arrays as well as genes were clustered using Wards linkage. Coloured bars at the top of the plot correspond to each treatment group and show good clustering of arrays based on treatment.

Table 8 Pairwise differential expression of GSE32138 RSV and H3N2 infected PBEC relative to time matched controls for gene in the PBEC - T cell activation gene set

	RSV		H3N2	
SYMBOL	log2FC	FDR.	log2FC	FDR.
IL6	1.21	1.42E-03	10.48	1.02E-14
IL1A	0.56	1.08E-02	8.87	2.89E-16
IL1B	0.15	7.01E-01	8.58	4.50E-14
TNF	0.36	8.24E-01	7.11	3.03E-14
IL23A	0.47	3.76E-02	7.09	7.08E-15
TSLP	0.06	9.03E-01	2.73	1.10E-07
IL33	0.25	6.49E-01	2.63	3.31E-06
LGALS9	2.38	6.38E-04	2.53	5.61E-05
IL12A	0.29	1.34E-01	2.01	2.09E-09
IL15	1.31	1.42E-07	1.74	5.33E-10
LGALS1	0.2	6.67E-01	1.6	7.58E-05
IL12B	-0.1	8.55E-01	1.53	1.44E-04
IL4	0	9.97E-01	1.09	9.88E-05
IL18	-0.02	9.48E-01	1.07	3.24E-05

	RSV		H3N2	
SYMBOL	log2FC	FDR.	log2FC	FDR.
CD274	1.9	3.94E-05	6.22	5.03E-12
CEACAM1	2.87	2.61E-08	5.74	3.46E-13
CD83	0.99	1.99E-02	5.14	2.59E-10
CD200	0.13	7.52E-01	3.4	3.13E-09
CD70	1.7	2.42E-06	3.22	1.22E-10
B2M	1.26	4.53E-05	2.91	3.88E-04
HLA-G	2.43	4.21E-03	2.57	1.45E-03
CD47	1.42	8.94E-03	2.04	8.10E-08
HLA-F	1.91	4.61E-09	1.84	1.79E-08
HLA-E	1.78	4.79E-05	1.44	1.50E-05
HLA-A	1.88	7.39E-09	1.43	2.08E-08
HLA-B	1.25	1.24E-07	1.18	2.57E-07
CD276	0.35	9.43E-01	1.14	4.02E-05
HLA-DPA1	1.24	1.18E-02	-0.68	4.81E-02

Log2 fold change (log2FC) values were calculated using pairwise differential expression analysis of RSV or H3N2 infected PBEC ALI relative to mock infected controls. Results are shown for genes mapped to the GO:BP term "T cell Activation" split into tables based on the function of the gene product. A) Secreted cytokines, B) surface bound eg receptors. Values are arranged by decreasing log2FC from the H3N2 data. Adjusted P values (FDR) are corrected for false discovery of multiple hypothesis testing using eBayes (limma). Values are coloured by significance in a 3-point scale with yellow indicating AdjP > 0.05.

The final PBEC-T Cell signature generated from the GSE32138 data set consisted of 34 genes with varied expression in response to RSV and H3N2. Almost half, 47% (16/34) were not DE by RSV, primarily genes from the secreted cytokine sub list. H3N2 also had minimal impact on 5 of these genes however the response to H3N2 was overall considerably stronger than RSV.

3.2.2 GSE47961 Analysis

GSE47961 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE47961) produced by Mitchell et al, looks at two highly pathogenic viruses, a Californian isolate of the 2009 H1N1 strain of IAV and SARS-CoV1, the causative agent of the 2002 SARS outbreak. In this dataset, samples were collected at multiple time points post infection allowing for analysis of temporal differences in the epithelial response to infection with these viruses.

Primary Human- tracheobronchial epithelial cells were differentiated using the ALI method. There is no specific description of donor allocation or number in the associated metadata, however the original authors state that the same cell stock was used for all replicate samples. Arrays for infected samples were performed in quadruplicate and mock-infected samples were analysed in triplicate, except for the 18h mock-infected array which was performed in duplicate. After ALI differentiation, the cells were exposed at the apical surface to either H1N1 (MOI-2) or SARS-CoV1 (MOI-2) for 2 hours before washing and incubation up to a maximum of 48h (H1N1) or 72h (SARS-CoV1).

This project re-analysed data from the 0, 6, 12, 18, 24, 36 and 48h H1N1 time points as well as the 0, 24, 48, 60, 72, 84 and 96h SARS-CoV1 time points in addition to the corresponding mock-infected samples. The mock-infected samples were not infection specific so where time points overlap the same mock samples were used. The final data import consisted of 88 Agilent single colour arrays, 56 infected and 32 mock infected, which as with the previous dataset, contain 45015 probes per array.

3.2.2.1 Pre-Processing and Quality Control testing.

The GSE47961 arrays were imported to R-Studio and normalised using the same approach described for the GSE32138 data. There were no substantial differences between the raw array intensity values and *NEQC* background correction and quantile normalisation resulted in appropriately distributed arrays suitable for DE analysis **Figure 3-10.**

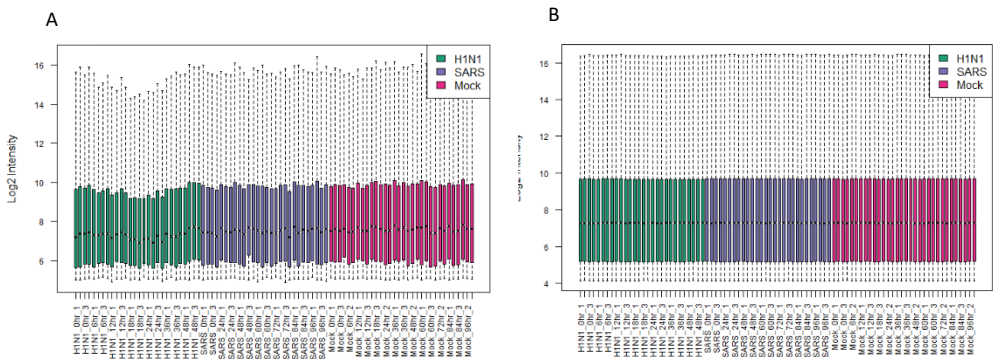


Figure 3-10 GSE47961 array QC: Boxplots demonstrate successful NEQC normalisation and background correction

GSE47961 array intensity values were normalised using the NEQC function from the limma package. This approach uses the control probe data included with the raw files to perform background correction followed by quantile normalisation. Intensity values for all arrays are visualised as boxplots both before (A) and after (B) normalisation. Arrays are coloured according to group membership.
NEQC normalisation resulted in evenly distributed intensity values suitable for differential expression analysis

No outliers were detected using MDS plots of normalised expression data and the arrays were clustered by treatment group overall **(Figure 3-11 A).** Treatment group clustering occurred along the first dimension (Dim 1) which explained 45% of the group variation. H1N1 arrays (Green) are most distinct while SARS-CoV1 arrays (Purple) are in close proximity to mock infected arrays (Pink) with some overlap of SARS-CoV1 24h arrays. All 0h samples predictably group with controls regardless of infection group. H1N1 arrays additionally separate along Dim1 by time point; 6hpi samples are closer to SARS-CoV1 and mock arrays followed by 12h and 48h samples, then 18 and 36h samples, with 24h samples forming a tight cluster furthest from the other arrays. The 18h and 36h H1N1 infections are less distinct and arguably form two mixed clusters separated along Dim2 (10% of variation).

SARS-CoV1 arrays also formed loosely time point dependent groups; the 96h arrays were clustered together, and separated from 48h arrays by Dim2, and 72h and 84h arrays were closer to H1N1 6h arrays. However, most replicates of the 60h, 72h and 84h time points did not cluster together.

Considerable overlap between SARS-CoV1 arrays and Mock-infected controls as well as clear separation of H1N1 arrays can be see with hierarchical clustering **Figure 3-11 B.**

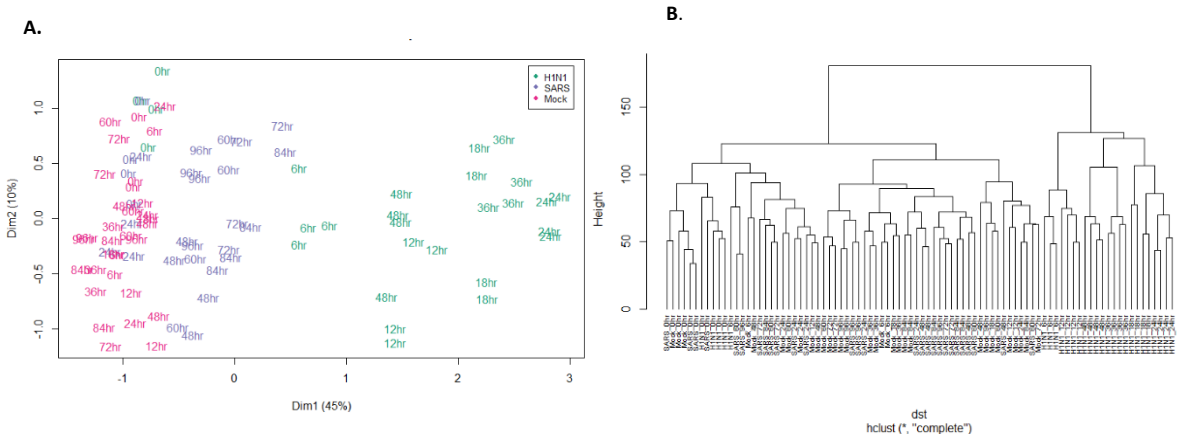


Figure 3-11 GSE47961 arrays cluster by treatment group using multidimensional scaling and Hierarchical clustering

A) MDS plot showing the variation between normalised intensity values from the GSE47961 data. Spots represent individual arrays coloured by viral agent (SARS-CoV1 = Purple, H1N1 = Green) or mock infected control (Pink) used during PBEC infection. The distance between points approximates the greatest log2FC between samples.
B: Hierarchical clustering of the same data using ward linkage.

3.2.2.2 H1N1 and SARS-CoV1 Differential expression analysis

The arrays from both infection groups were normalised together and treated as a single dataset for time matched DE analysis. There were no condition specific mock-infected controls in this dataset so where time-points overlap, the H1N1 and SARS-CoV1 infection samples were analysed relative to the same baseline sample. Time matched DE analysis identified both upregulated and downregulated genes at every time point post H1N1 infection (**Table 9)**. As early as 6hpi there are over 200 upregulated genes and the number of genes increases with increasing time point until 48hpi when there is a relative drop in DE numbers. The highest number of genes were DE at 24hpi and 36hpi with a roughly even number of genes upregulated and downregulated beyond the significance and effect size thresholds at these time points.

SARS-CoV1 infection had a considerably weaker impact on gene expression than H1N1, the greatest amount of total DE occurred at 48hpi with genes evenly up and downregulated. Slightly more genes were upregulated at 72 and 84hpi however there were still fewer genes DE by SARS-CoV1 at these time points than there were at 6hpi with H1N1. DE gene counts at these time points are lower than the number of genes DE by H1N1 at 6hpi. No time points besides 48h resulted in downregulation with the exception of 3 genes downregulated at 24hpi and 1 downregulated at 96hpi.

Table 9 Differentially expressed gene counts for H1N1 and SARS-CoV1 infection relative to time matched controls

H1N1			
HPI	**NotSig**	**UP**	**Down**
6	18561	248	3
12	18115	580	117
18	17405	949	458
24	16284	1231	1297
36	16618	1195	999
48	17603	565	644

SARS-CoV1			
HPI	**NotSig**	**UP**	**Down**
24	18798	11	3
48	18575	111	126
60	18715	91	6
72	18648	164	0
84	18645	167	0
96	18714	97	1

Differential expression analysis performed on NEQC normalised expression data. HPI = Hours Post Infection. Genes are counted as differentially expressed with a log2FC ≤ -1 or ≥1 AND an adjusted P value < 0.05. NotSig includes all genes not meeting both criteria. Expression data has been collapsed to gene level by taking the average of probe variants.

As multiple timepoints were available with this data set it was possible to gain a deeper understanding of the impact of these viruses over the course of infection. To do this, the maSigPro package was used to analyse normalised expression data from each infection as individual time series with the results described separately for the H1N1 and SARS-CoV1 data.

3.2.2.2.1 H1N1 time series analysis

H1N1 infection resulted in significant DE of genes from the earliest 6h time point with both up and downregulated genes showing peak DE at 18-24hpi. The maSigPro clustering approach identified 9 distinct temporal expression patterns for a total of 7317 genes classified as DE in at least one time point relative to control samples. Plots showing the trend for expression of genes in each cluster are shown in **Figure 3-12.** Most clusters of genes show distinct expression to control samples; clusters 5, 7 and 8 are upregulated while clusters 1,2,3,4 and 9 show overall downregulation relative to uninfected samples. The genes in cluster 6 are also ultimately upregulated by infection following 12 hours of relative inhibition. In most clusters, the 24h time point shows the greatest difference in expression between H1N1 infected and control samples.

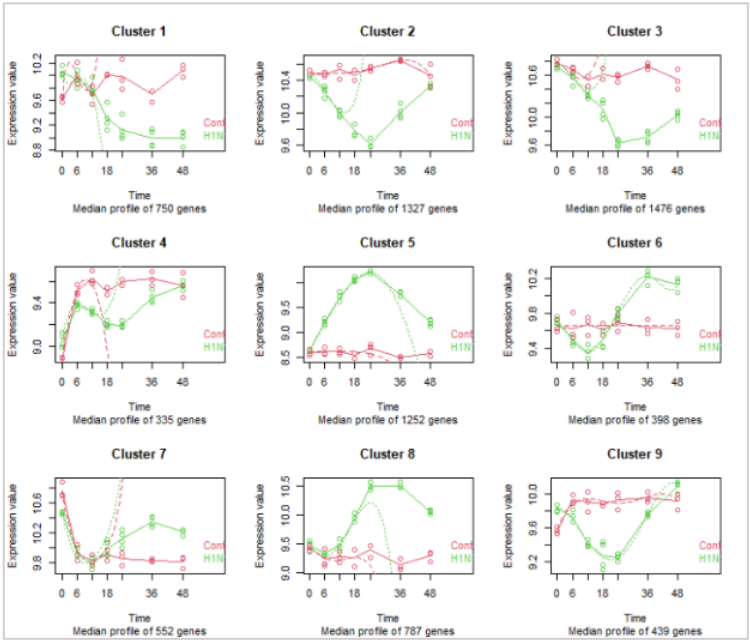


Figure 3-12 GSE4791 H1N1 Time series analysis identified clear distinction between infected and control samples over the course of infection

NEQC normalised array intensity data for H1N1 infected samples were analysed a complete time series using maSigPro. Genes with significant differences between experimental conditions are clustered based on shared expression patterns over the course of infection. Lines show the average expression of genes within that cluster. Expression in control samples are coloured Red and infected samples are coloured Green. The number of genes within each cluster is include below each plot.

As the expression values for each cluster peak at 24hpi, log2FC values from the 24h time-matched DE analysis were used to rank genes from each cluster for ORA using clusterProfiler. Genes not significantly DE at this time point were excluded from the analysis and the complete list of ranked genes DE at 24hpi were used as the background or "universe" gene set.

The result of ORA of the genes downregulated by H1N1 infection, clusters 1, 2, 3 and 4, are shown in **Figure 3-13.** Cluster 9 was not included as none of the 439 genes in this group were DE at the 24h peak relative to time-matched controls. There were only 18/335 genes from cluster 4 with log2FC exceeding -1 at the same time point, ORA was performed for these genes however no GO:BP terms were significant. The downregulated genes with both cluster 2 and cluster 3 expression patterns were significantly enriched for pathways relating to cellular metabolism. For example, the term "regulation of cellular metabolic processes" (GO: 0031323) was significant for cluster 2 with 121 genes mapped to this term, while 100 genes from the term "catabolic process" (GO: GO:0009056) were found in cluster 3. Both cluster 2 and 3 also share similar expression patterns with the extent of downregulation increasing until 24hpi and then decreasing again until the 48 end of infection. The expression values of cluster 2 genes are similar in both infected and control samples by 48hpi whereas cluster 3 genes remain downregulated by infection, relative to control samples, throughout. Cluster 1 genes were primarily associated with GO:BP terms relating to cilia, the top 10 most enriched terms for this gene list included "cilium organisation" (GO:0044782) and "microtubule-based movement" (GO:0007018). Expression of these genes declines from 12hpi to a stable low relative to controls at 24hpi. Their expression remains at this level until the 48h endpoint.

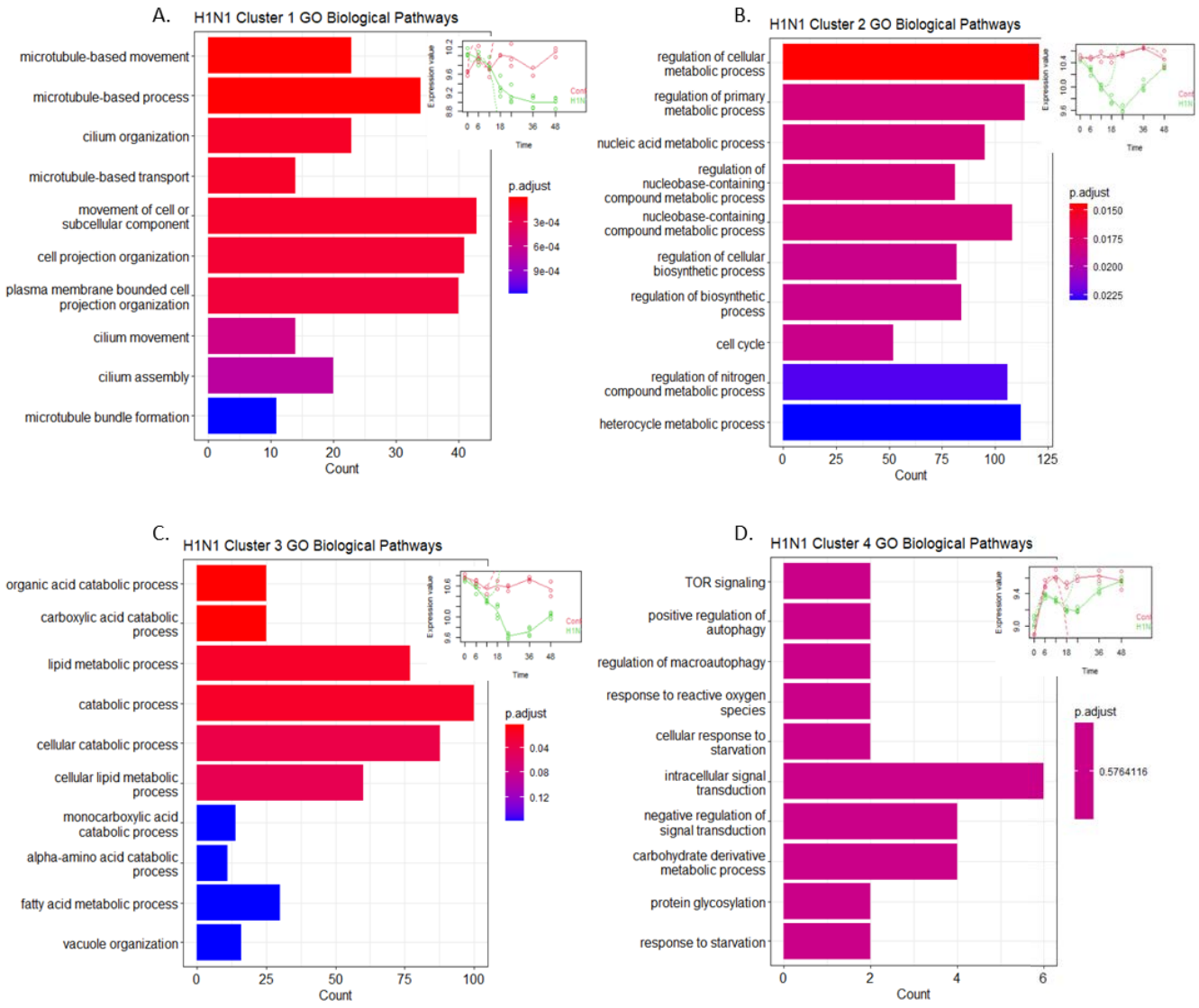


Figure 3-13 GSE47961 H1N1 downregulated gene cluster ORA shows a focus on cellular metabolism and cilia organisation

Over representation analysis (ORA) of significantly DE genes from each downregulated gene cluster was performed using the full list of downregulated genes from 24h time-matched analysis as the background gene list. The 10 most significantly enriched GO:BP terms are shown as bar plots with the length of each bar indicating the number of genes mapped to each term. Bars are coloured by significance as described in each key. Plots were generated using the clusterprofiler R package following ORA with enrichGO. Genes from clusters 1, 2, 3 and 4 are shown in plots A, B, C and D respectively. The expression pattern of genes in each group, determined using maSigPro, are inset with each bar plot.

Genes upregulated by H1N1, from clusters 5, 6, 7 and 8, were analysed using the same approach with all significant DE genes, with log2FC >= 1, from the 24hpi time-matched analysis used as background genes. No GO:BP terms were significantly enriched for the cluster 6 gene list **(Figure 3-14 B)**, however the few significantly DE genes in this cluster were mapped to terms relating to cellular metabolism. Genes in cluster 6 show a temporal expression pattern with initial downregulation followed by upregulation between 18 and 36h. Cluster 5 genes are rapidly and increasingly upregulated between 6 and 24hpi followed a reduction in the extent of upregulation until 48h, approaching levels seen at 6hpi by the end of infection. The genes in this cluster were highly significantly enriched for terms such as “response to virus” (GO:0009615) and “regulation of viral genome replication” (GO:0045069) (**Figure 3-14 A.).**

Genes grouped into cluster 8 begin being upregulated by H1N1 infection at 12-18hpi and their expression value increases to a plateau between 24 and 36hpi before decreasing to levels comparable with the 18h time point by 48hpi. While not reaching the levels of significance seen for cluster 5, GO:BP terms relating to adaptive immunity such as “T cell mediated immunity” (GO:0002456) were significantly represented the cluster 8 gene list **(Figure 3-14 D).**

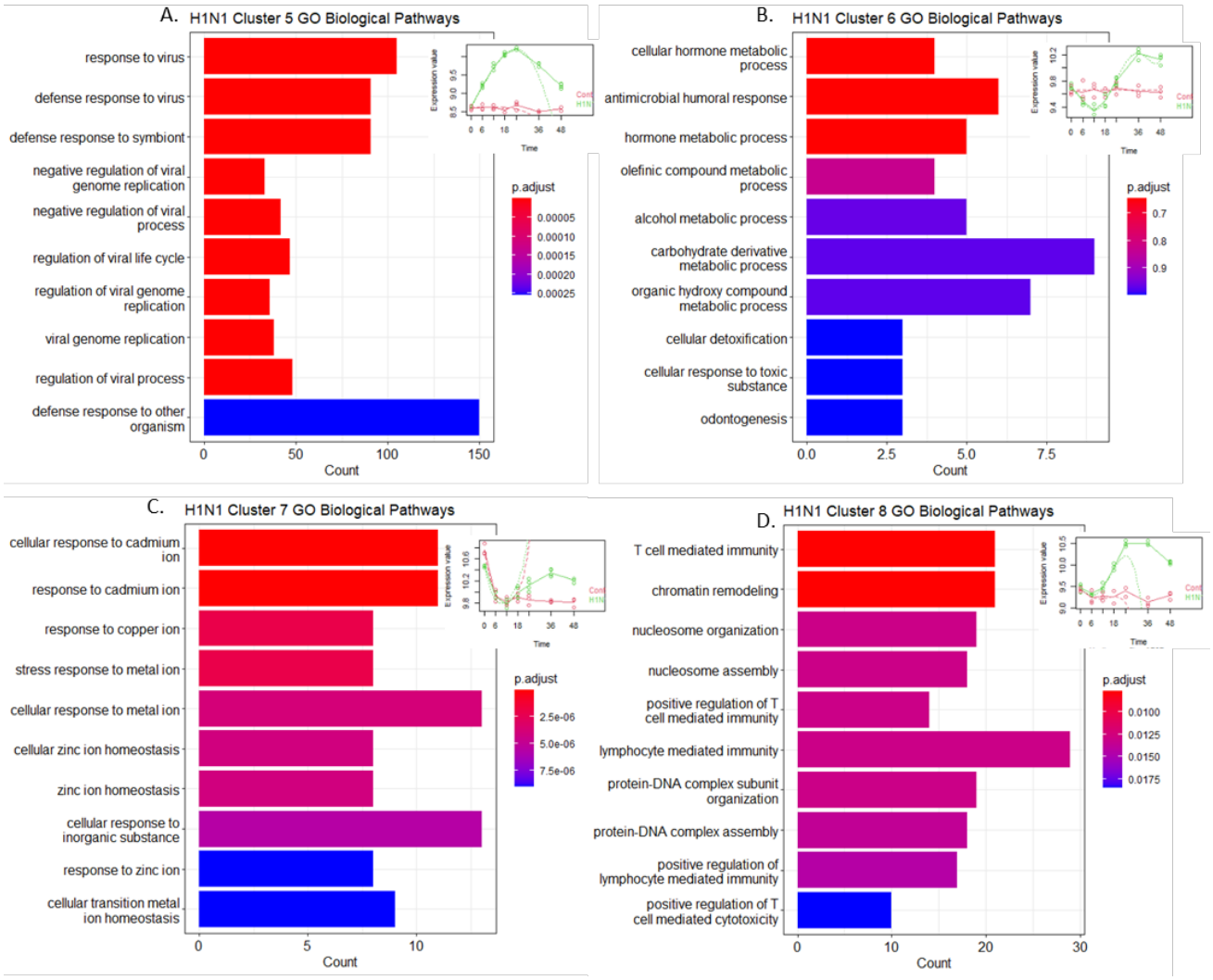


Figure 3-14 GSE47961 H1N1 upregulated genes were enriched from immune, antiviral and metabolic pathways

Over representation analysis (ORA) of significantly DE genes from each upregulated gene cluster was performed using the full list of upregulated genes from 24h time-matched analysis as the background gene list. The 10 most significantly enriched GO:BP terms are shown as barplots with the length of each bar indicating the number of genes mapped to each term. Bars are coloured by significance as described in each key. Plots were generated using the clusterprofiler R package following ORA with enrichGO. Genes from clusters 5, 6, 7 and 8 are shown in plots A, B, C and D respectively. The expression pattern of genes in each group, determined using maSigPro,are inset with each barplot.

Antiviral gene expression predominantly took the immediate cluster 5 pattern, however there were also genes in cluster 8 which mapped to terms such as “defence response”. Normalised expression values for genes from both clusters mapped to this term are shown in **Figure 3-15**. IFN inducible genes (eg. *IFIT1, IFIT2, ISG15, OAS1, OAS2, IDO1*) as well as the type I *(IFNB1* and *IFNA14)* and type III (*IFNL1*, *IFNL2*, *IFNL3*) IFN genes were found in cluster 5. As with H3N2, H1N1 infection had a stronger

impact on type III IFNs than type I IFNs especially at the 24hpi peak of expression. A table showing log2FC values for key antiviral genes at all time points is included (**Table 10**).

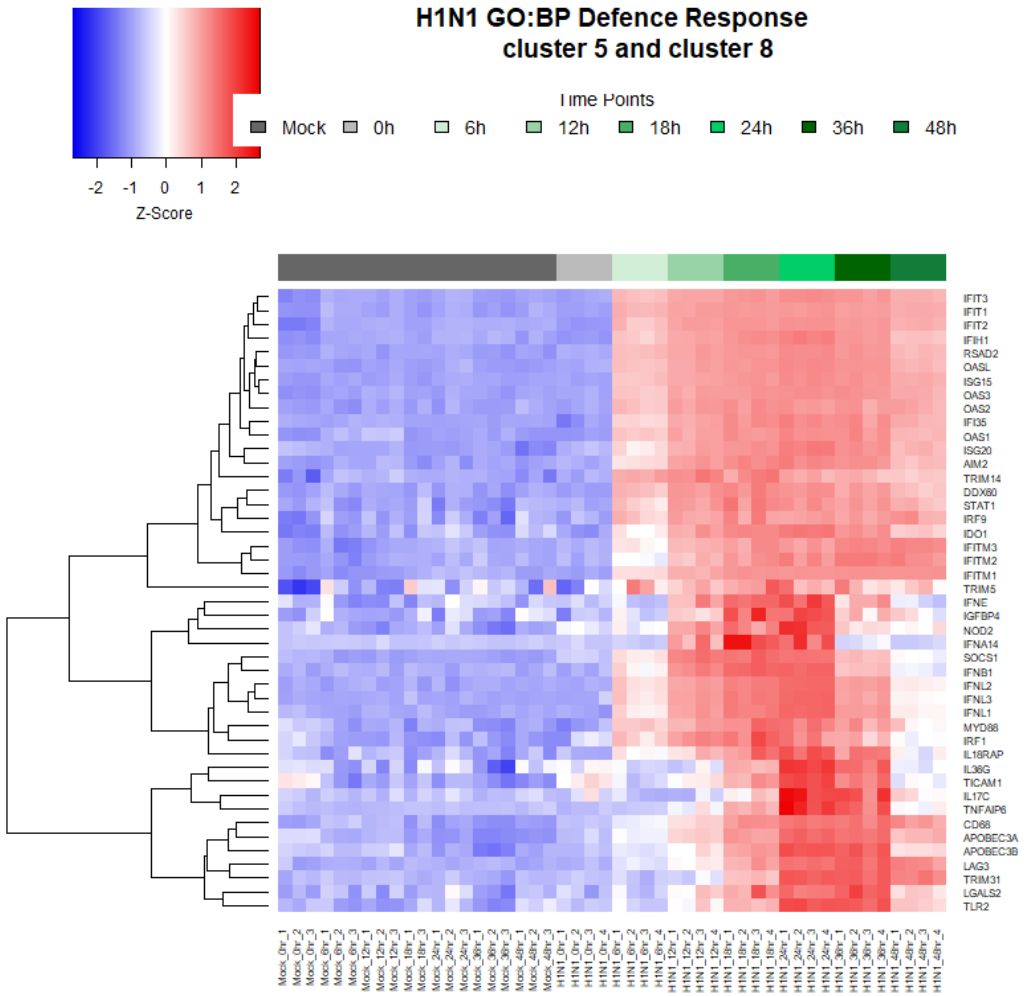


Figure 3-15 GSE47961 H1N1: had a strong and immediate impact on many genes from the term Defence Response

GSE47961 H1N1 expression data for genes matched to the GO:Biological process "Defence Response" at all time points.. A heatmap was generated from neqc normalised intensity values for arrays from all H1N1 infected and mock-infected time points. Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Blue shading indicates that expression is at or greater than 2 SD lower than the mean for a given gene and red shading indicates genes are expressed at or greater than 2 SD above the mean. Group membership of H1N1 or mock infected arrays is shown by the coloured column bars and key.

Table 10 Time matched Differential expression analysis of GSE47961 H1N1 infected samples confirms the strong induction of antiviral genes mapped to the term "Defence Response".

	H1N1 6h		H1N1 12h		H1N1 18h		H1N1 24h		H1N1 36h		H1N1 48h	
SYMBOL	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**
IFNL3	5.228	3.70E-23	7.567	3.80E-32	8.427	7.96E-32	9.777	9.17E-39	7.547	1.92E-32	4.119	2.20E-18
IFNL2	5.292	8.35E-24	7.421	4.19E-32	8.778	4.19E-33	9.716	4.84E-39	7.447	1.59E-32	4.293	1.45E-19
OASL	7.590	4.01E-34	8.453	1.83E-37	9.391	5.54E-37	9.152	5.90E-40	9.490	1.96E-40	8.022	5.72E-36
IFNL1	4.425	6.59E-22	6.563	2.13E-31	7.759	1.78E-32	8.996	1.90E-39	7.117	1.12E-33	3.850	3.00E-19
RSAD2	6.292	6.81E-31	7.591	4.11E-36	7.994	1.16E-33	8.287	2.06E-38	8.399	1.06E-38	6.754	3.70E-33
IFIT2	6.018	3.07E-33	7.469	1.28E-38	8.390	3.13E-38	7.957	1.53E-40	7.659	8.99E-40	6.596	2.59E-36
IFIT1	6.083	1.47E-32	7.339	1.78E-37	7.990	9.85E-37	7.453	1.29E-38	7.868	1.14E-39	6.730	1.91E-35
IFIT3	5.491	1.51E-32	6.586	1.83E-37	7.452	2.78E-37	7.220	3.67E-40	7.534	9.15E-41	6.297	5.17E-36
IFNB1	2.422	1.11E-09	5.200	1.74E-24	5.822	2.80E-24	6.091	1.61E-28	3.622	2.69E-17	1.651	4.93E-06
OAS3	4.334	3.99E-26	5.114	8.07E-31	5.890	1.46E-31	5.828	6.11E-34	5.930	2.45E-34	5.062	4.01E-30
IDO1	3.382	6.57E-10	5.368	2.49E-18	6.467	1.09E-19	5.752	3.42E-20	6.762	1.45E-23	3.190	1.08E-09
OAS1	3.537	2.08E-24	3.969	1.72E-27	5.688	5.42E-33	5.628	3.77E-36	5.378	5.06E-35	4.782	1.24E-31
ISG15	4.121	1.65E-18	4.946	5.33E-22	5.346	1.67E-19	5.330	3.96E-23	5.066	2.69E-20	4.417	1.23E-16
OAS2	4.149	1.40E-23	4.869	8.05E-27	5.487	7.15E-26	5.145	2.13E-28	5.504	7.52E-29	4.444	2.24E-24
IL36G	1.029	3.35E-01	0.789	2.76E-01	3.119	1.08E-05	5.038	2.20E-13	6.544	5.47E-18	0.169	8.73E-01
IFIH1	3.107	6.20E-31	3.808	2.58E-36	4.450	5.54E-37	4.433	2.92E-40	3.981	1.06E-37	3.401	2.90E-33
STAT1	3.287	2.90E-14	4.011	7.39E-20	4.100	3.73E-18	4.272	7.84E-23	4.464	1.48E-23	3.131	1.57E-14
ISG20	2.160	7.61E-20	3.179	6.40E-29	3.861	1.26E-30	4.203	6.03E-36	3.708	7.25E-33	3.638	7.35E-32
LAG3	0.985	3.75E-05	1.787	4.17E-13	2.922	3.68E-20	4.016	1.72E-30	4.136	3.98E-31	3.654	1.14E-27
IL17C	-0.006	9.98E-01	0.025	9.62E-01	1.329	1.73E-04	3.630	6.26E-20	3.399	1.44E-18	1.041	9.34E-04
SOCS1	1.461	1.83E-01	3.582	1.92E-17	3.746	8.04E-17	3.510	4.05E-17	2.390	1.90E-09	1.121	1.52E-01
IFITM1	2.694	4.16E-27	3.529	7.02E-34	3.649	8.06E-32	3.283	1.36E-32	3.507	4.22E-34	3.878	6.88E-36
TLR2	0.636	1.88E-02	1.586	4.31E-11	1.817	1.86E-11	3.094	3.05E-24	3.737	1.68E-28	1.151	2.85E-07
IFITM2	1.357	1.02E-08	2.116	2.75E-16	2.542	1.57E-17	2.657	2.63E-21	2.814	1.85E-22	2.703	4.12E-21
IRF1	1.315	9.99E-07	1.951	1.68E-12	2.445	2.38E-14	2.062	5.42E-14	1.603	3.70E-10	0.941	1.32E-04
MYD88	1.274	4.27E-07	1.714	2.83E-13	2.367	3.44E-17	2.018	7.81E-17	2.047	4.99E-16	1.114	1.69E-07
NOD2	0.624	8.17E-05	1.602	8.26E-19	1.546	8.03E-16	2.013	4.30E-24	1.329	1.10E-15	0.791	9.01E-08
IFNA14	0.025	9.76E-01	1.383	3.18E-15	2.467	1.15E-24	2.003	3.98E-23	0.281	5.54E-02	0.050	8.38E-01
IRF9	1.457	2.73E-09	1.837	2.95E-13	1.894	6.52E-12	1.919	1.45E-14	2.342	3.77E-18	1.671	8.08E-12
IL18RAP	0.747	1.20E-03	0.969	3.54E-06	1.877	3.53E-13	1.861	1.43E-15	1.930	3.65E-16	0.679	7.96E-04
IFNE	0.221	7.39E-01	1.670	1.53E-10	2.164	8.57E-13	1.853	1.02E-12	1.161	8.14E-07	0.283	3.22E-01

GSE47961 H1N1 expression data was normalised using *NEQC* and differentially expressed (DE) genes were identified using LIMMA, with time matched mock-infected samples used as baseline. The result of this analysis for genes mapped to the GO:BP term "Defence Response" are shown as log2 fold change (log2FC) and FDR adjusted P values (FDR). Genes are arranged by log2FC at 24hpi and FDR values are shaded with a 3 point scale from yellow – green – violet, indicating increasing significance.

3.2.2.2.2 SARS-CoV1 Time series analysis

In contrast to H1N1, expression patterns of SARS-CoV1 infected samples were very similar to the control samples for all but clusters 4, 6 and 9. Genes grouped into cluster 1 change expression throughout infection but to the same extent in both infected and control samples. Cluster 4 has the most definitive SARS-CoV1 infection dependent upregulation with stable expression in control samples. The upregulation seen for genes in cluster 9 is more modest and occurs beyond control samples only at 60hpi. Cluster 6 appears to have relatively stable high expression compared to the uninfected controls at all time points. Plots generated by maSigPro showing the result of this analysis are shown in **Figure 3-16.**

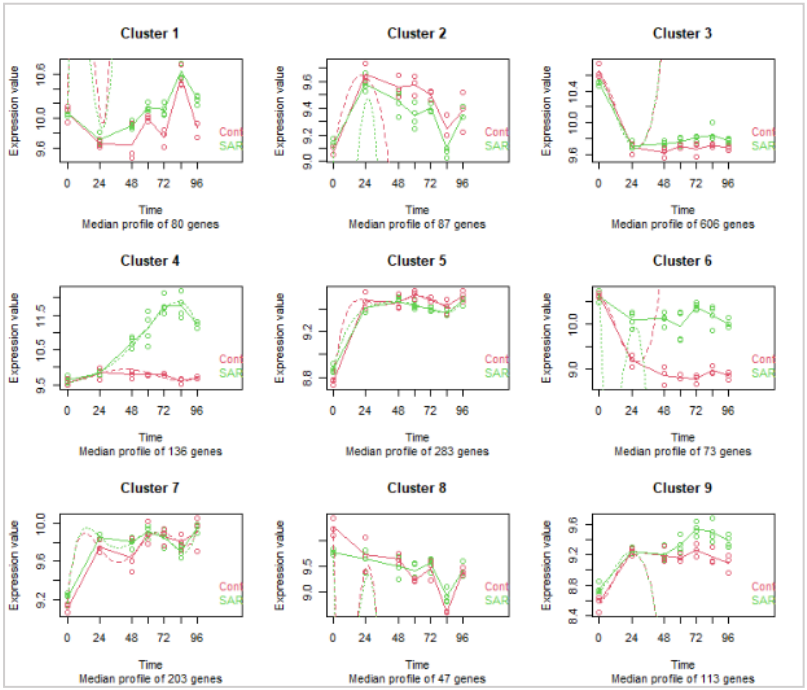


Figure 3-16 GSE47961: SARS-CoV1 Time Series analysis shows very few genes with differences in temporal expression pattern between infected and control samples

NEQC normalised array intensity data for SARS-C0v1 infected samples were analysed as a complete time series using the maSigPro package. Genes with significant differences between experimental conditions are clustered based on shared expression patterns over the course of infection. Lines show the average expression of genes within that cluster. Expression in control samples are coloured Red and infected samples are coloured Green. The number of genes within each cluster are included below each plot.

ORA of genes in cluster 4 and cluster 6 was performed using all SARS-CoV1 72hpi upregulated genes as background. Cluster 9 was also explored however there were only 2 genes with log2FC >= 1 in this cluster. As genes from the remaining clusters did not differ from control samples they were not investigated further. The 10 most significantly enriched terms from cluster 6 ORA **(Figure 3-17. B)** included "cellular response to stress" (GO:0033554) and "developmental process" (GO:0032502). There was considerable overlap between these terms with transcription factor encoding genes such as *FOS*, and *Early Growth Response 1* (*EGR1*) found in all 10 shown here. Genes in this cluster were consistently expressed with values greater than those for control samples however, there was more definitive upregulation of genes grouped into Cluster 4. These genes are steadily upregulated to a peak of expression 72h post SARS-CoV1 infection. The genes in cluster 4 have a clear antiviral focus with "response to virus" and "innate immune response" among the 10 most significant terms for this cluster of genes **(Figure 3-17 A).**

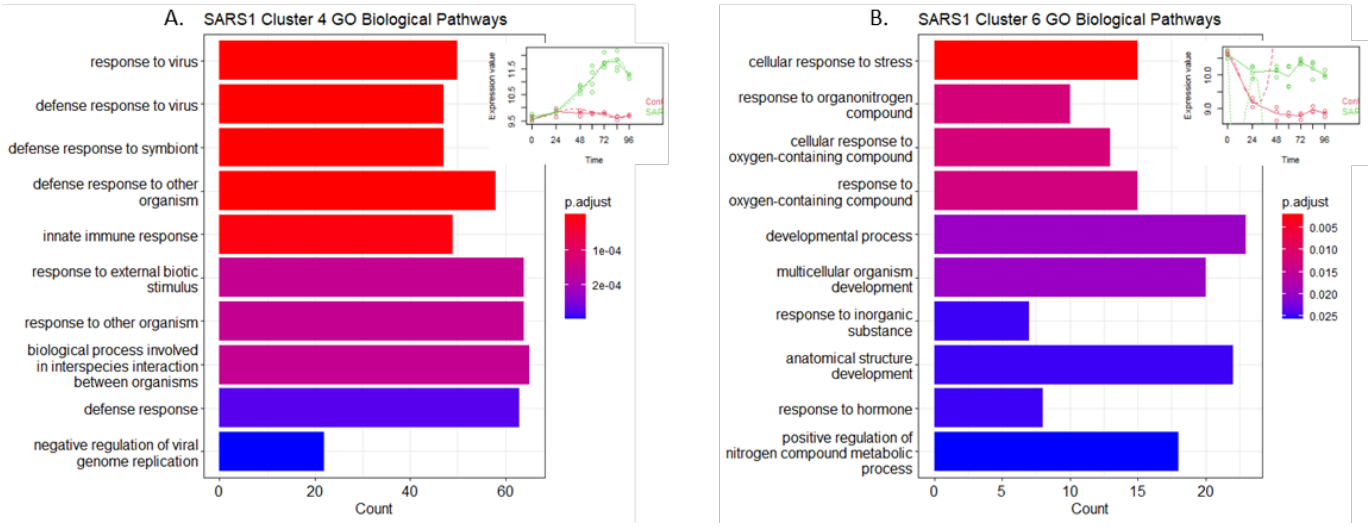


Figure 3-17 GSE47961 SARS-CoV1 upregulated genes have cluster specific roles in the antiviral and developmental responses

Over representation analysis (ORA) of significantly DE genes from each upregulated gene cluster was performed using the full list of upregulated genes from 72h time-matched analysis as the background gene list. The 10 most significantly enriched GO:BP terms are shown as bar plots with the length of each bar indicating the number of genes mapped to each term. Bars are coloured by significance as described in each key. Plots were generated using the clusterprofiler R package following ORA with enrichGO. Genes from clusters 1, 2, 3 and 4 are shown in plots A, B, C and D respectively. The expression pattern of genes in each group, determined using maSigPro, are inset with each bar plot.

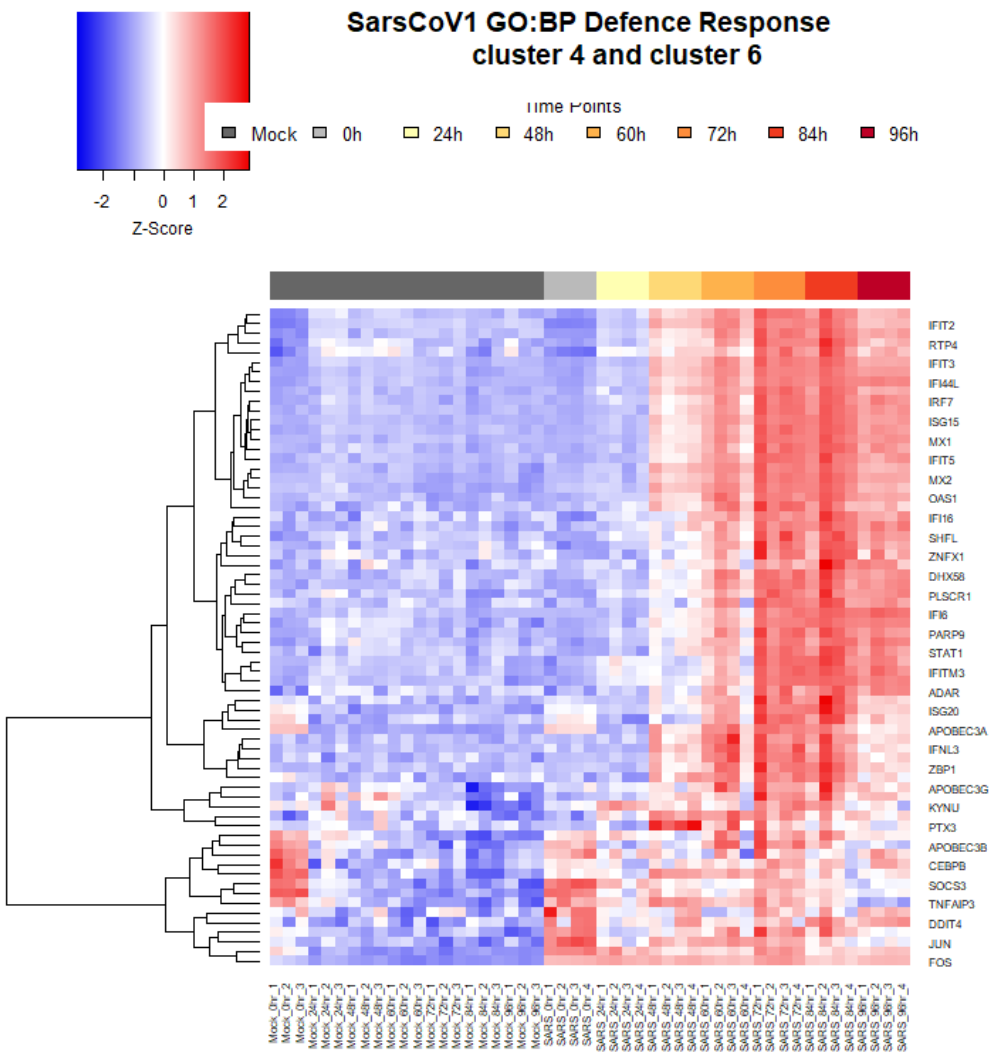


Figure 3-18 GSE47961 SARS-CoV1 infection resulted in delayed induction of antiviral genes mapped to the term defence response

GSE47961 SARS-CoV1 expression data for genes with the cluster 4 or cluster 6 expression pattern over 72h infection matched to the GO: Biological process "Defence Response" at all time points. Genes in these clusters were annotated with GO:BP terms due to their clear upregulation relative to controls in maSigPro analysis A heatmap was generated for these genes from neqc normalised intensity values for arrays from all SARS-CoV1 infected and mock-infected time points.
Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Strong Blue shading indicates that expression is at or greater than -2 SD lower than the mean for a given gene and red shading indicates genes are expressed at or greater than 23 SD above the mean. Group membership of SARS-CoV1 time points or mock infected arrays is shown by the coloured column bars and key.

Genes from cluster 4 including antiviral IFIT (inc. *IFIT1*, *IFIT3*), IRF (*IRF7*, *IRF9*), ISG (*ISG15*, *ISG20*) and OAS (*OAS1, OAS2, OAS3, OASL*) families of genes, begin being upregulated by SARS-CoV1 at 48h and reach a plateaux of expression between 72 and 84hpi. The genes mentioned here were upregulated with log2FC between 3 and 6 at 72hpi. Type I and type III IFN genes, *IFNB1*, *IFNL1*, *IFNL2* and *IFNL3,* were also grouped into cluster 4 and were all roughly equally upregulated (log2FC 1-2) by SARS-CoV1

at the 72h peak of expression. “Defence response” genes were also present in cluster 6, with less significant enrichment. Normalised expression values for genes mapped to “defence response” from both gene clusters are shown as a heatmap **(Figure 3-18),** and log2FC **(Table 11).**

Table 11 Time matched Differential expression analysis of GSE47961 SARS-CoV1 infected samples confirms the later induction of antiviral genes mapped to the term "Defence Response".

	SARS1 24h		SARS1 48h		SARS1 60h		SARS1 72h		SARS1 84h		SARS1 96h	
SYMBOL	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**
OASL	-0.238	9.37E-01	3.857	5.24E-18	4.933	1.66E-23	5.984	1.70E-28	6.168	3.83E-29	4.548	7.02E-22
IFI44L	-0.006	1.00E+00	2.875	3.00E-14	3.902	1.61E-20	5.531	7.85E-29	5.825	8.55E-30	5.514	3.27E-28
IFIT1	0.266	8.79E-01	3.464	7.67E-19	4.229	1.88E-23	5.312	4.02E-29	6.246	1.47E-32	4.687	6.96E-26
RSAD2	-0.164	9.52E-01	3.029	4.50E-15	3.463	4.49E-18	4.588	1.61E-24	4.755	6.63E-25	3.445	2.31E-17
IFI27	-0.700	5.99E-01	0.920	9.63E-02	2.350	3.68E-06	4.490	6.37E-16	5.135	4.91E-18	4.642	2.20E-15
IFIT2	-0.303	7.84E-01	2.923	7.95E-17	3.648	7.34E-21	4.360	2.47E-25	4.602	1.81E-27	2.719	6.31E-16
OAS3	-0.037	9.97E-01	1.617	2.67E-07	2.505	5.36E-14	4.022	1.36E-23	3.815	3.14E-23	3.214	5.16E-19
ISG15	0.066	9.47E-01	1.873	3.77E-05	2.684	9.22E-08	3.531	4.56E-12	3.799	7.82E-15	2.793	1.37E-06
OAS1	-0.012	9.99E-01	1.510	5.45E-08	2.350	3.96E-15	3.282	1.98E-22	3.021	1.24E-20	2.571	4.62E-17
IFITM1	0.227	7.01E-01	0.901	1.58E-06	1.842	6.16E-18	3.087	1.10E-29	2.892	2.54E-28	3.454	5.17E-32
IRF7	0.040	9.97E-01	0.915	3.00E-02	1.564	1.10E-03	2.639	1.00E-10	2.578	1.91E-10	2.077	6.01E-07
IFNL2	0.118	9.94E-01	1.194	7.77E-03	2.580	1.42E-09	2.525	1.42E-09	2.273	3.82E-08	1.385	2.93E-03
IFNL3	-0.231	9.58E-01	1.471	1.03E-03	2.060	1.69E-06	2.390	1.31E-08	1.894	7.17E-06	1.040	6.53E-02
IFNB1	0.128	9.93E-01	1.379	1.31E-03	2.419	4.40E-09	2.384	3.40E-09	2.573	2.96E-10	1.159	1.80E-02
IFNL1	-0.032	9.99E-01	1.304	1.16E-03	1.544	9.82E-05	2.080	3.05E-08	1.861	7.17E-07	0.467	5.46E-01
IFITM3	0.281	6.96E-01	0.324	2.38E-01	0.952	3.84E-05	2.006	3.23E-16	2.174	8.65E-18	2.160	2.13E-17
ISG20	0.011	9.99E-01	0.962	4.34E-06	1.261	1.57E-09	1.926	3.44E-17	2.044	2.20E-18	1.393	3.95E-11
IRF9	-0.124	9.68E-01	0.191	6.14E-01	0.961	3.16E-04	1.743	1.12E-11	1.704	2.76E-11	1.722	2.79E-11

GSE47961 SARS-CoV1 expression data was normalised using NEQC and differentially expressed (DE) genes were identified using LIMMA, with time matched mock-infected samples used as baseline. The result of this analysis for genes mapped to the GO:BP term “Defence Response” are shown as log2 fold change (log2FC) and FDR adjusted P values (FDR). Genes are arranged by log2FC at 72hpi and FDR values are shaded with a 3 point scale from yellow – green – violet, indicating increasing significance: Yellow > 0.05, Green: 0.05 – 1E-10, Purple: 1E-10 – 1E-40

While there were no clusters of genes with a pattern of downregulation identified using maSigPro, the 126 genes found to be downregulated by SARS-CoV1 infection at 48h using limma analysis were explored further. To perform ORA on this gene set, the full list of genes significantly DE by SARS-COV1 at 48h were used as the background gene set, without effect size filtering. The pathway terms identified for these genes related to extracellular matrix organisation and tissue development (**Figure 3-19).**

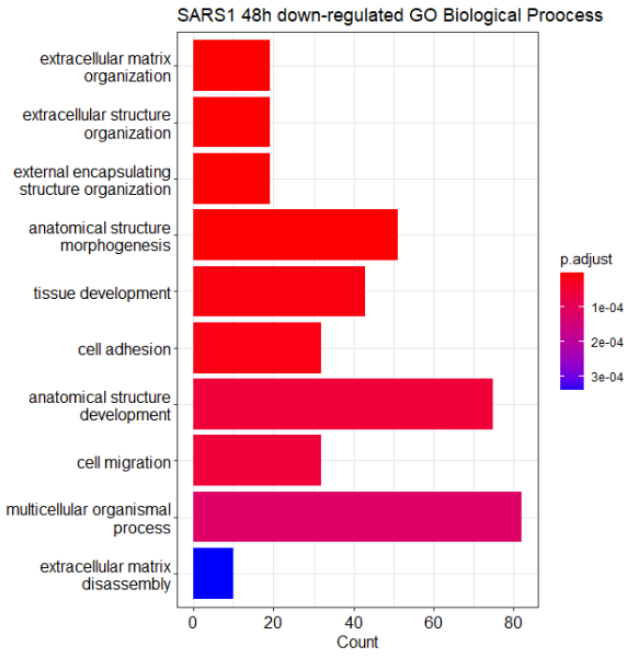


Figure 3-19 GSE47961 SARS-CoV1 downregulated genes were enriched for terms relating to the extracellular matrix.

Over representation analysis (ORA) of significantly downregulated genes at 48h post SARS-CoV1 infection was performed using the full list of downregulated genes from 48h time-matched analysis as the background gene list. The 10 most significantly enriched GO:BP terms are shown with the length of each bar indicating the number of genes mapped to each term. Bars are coloured by significance as described in the key. Plots were generated using the clusterprofiler R package following ORA with enrichGO.

3.2.2.3 H1N1 and SARS-CoV1 PBEC – T cell gene Expression

T cell related GO:BP term enrichment was explored using all genes which were upregulated at the time points with maximal DE. For H1N1 this was the 24h time point and for SARS-CoV1 the genes DE at 72hpi were combined with 38 which were uniquely DE at 84hpi. ORA analysis was performed for log2FC ranked up regulated genes from each infection separately using the full list of significantly (FDR < 0.05) DE genes without effect size filtering for background genes. From 1231 and 202 genes upregulated by H1N1 and SARS-CoV1 infection respectively, *clusterProfiler* mapped a total of 1088 and 190 genes to GO:BP pathway terms. The result of each ORA was filtered for terms relating to T cell recruitment or activation, without limiting results by significance of enrichment. This identified genes DE by each infection with roles in these functional terms. H1N1 infection caused DE of a greater number of genes within these terms than SARS-CoV1, with many also considered significantly enriched **Figure 3-20**. SARS-CoV1 infection resulted in upregulation of genes with T cell

related functions but to a lesser extent with no terms reaching significance**.** A direct comparison of gene counts per T cell term as well as FDR adjusted P values for enrichment are shown in **Table 12.**

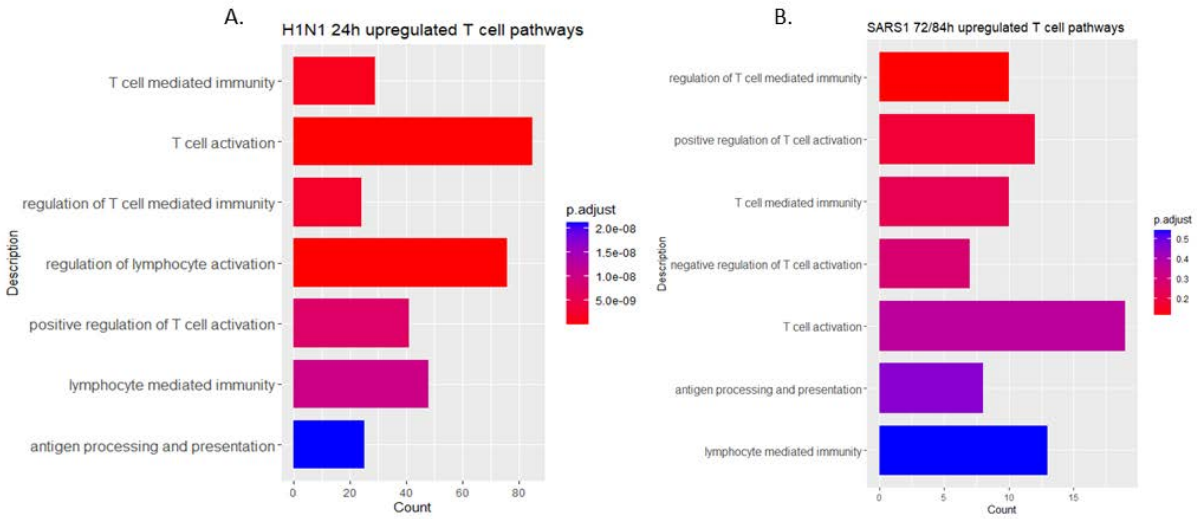


Figure 3-20 GSE47961 H1N1 infection resulted in differential expression of a greater numbers of T cell related genes than SARS-CoV1

Differentially expressed genes identified from the GSE47961 data set using limma, were annotated using over representation analysis with the clusterProfiler package. The 1231 genes significantly upregulated by H1N1 at 24hpi, and 202 genes significantly upregulated by SARS-CoV1 (SARS-1) at 72h or 84hpi, were analysed relative to all genes with adjusted p values < 0.05 in each infection analysis. The resulting GO biological process pathway terms identified for each virus were filtered to show pre-determined T cell activation or recruitment terms if they were enriched for each gene list. H1N1 infection (A) produced significant results for all terms. SARS-1 infection (B) was less significant.

Table 12 GSE47961 H1N1 and SARS-CoV1 T cell activation and recruitment GO BP term enrichment.

		H1N1	**(1088 genes)**	**SARS1**	**(190 genes)**
ID	**Description**	**Count**	**FDR**	**Count**	**FDR**
GO:0042110	T cell activation	85	1.99E-17	19	0.367
GO:0051249	regulation of lymphocyte activation	76	2.01E-12	20	0.268
GO:0002456	T cell mediated immunity	29	1.25E-09	10	0.227
GO:0050870	positive regulation of T cell activation	41	7.48E-09	12	0.182
GO:0002449	lymphocyte mediated immunity	48	1.03E-08	13	0.542
GO:0019882	antigen processing and presentation	25	2.12E-08	8	0.460
GO:0002706	regulation of lymphocyte mediated immunity	33	1.30E-07	11	0.179
GO:0072678	T cell migration	19	2.39E-07	0	na
GO:0050868	negative regulation of T cell activation	18	0.00015	7	0.285

T cell recruitment or activation related Gene ontology Biological process terms were identified from ORA performed for genes upregulated by H1N1 and SARS-CoV1 (SARS-1) using the clusterProfiler R package. The number of genes mapped to each term are shown along with the enrichment significance p value, adjusted for multiple hypothesis testing (FDR)

The 19 genes from the H1N1 gene list which mapped to “T cell migration” were reduced to 7 with relevance to PBEC recruitment of T cells. Expression values for these genes over 48h H1N1 infection are shown as a heatmap **(Figure 3-21).** Row clustering, with time point ordered columns, identified clear distinctions between genes with different temporal profiles. *CXCL10*, *CXCL11* and *CCL5* were clustered together with consistent low expression in all control samples and shared immediate upregulation and 24h peak expression in infected samples. These genes were grouped into cluster 5 by maSigPro and were among the most highly DE in time matched analysis. *CXCL10* and *CXCL11* were particularly highly upregulated with log2FC > 11 at 24hpi **(Table 13).** *CCL2* was also expressed with the cluster 5 pattern but with a smaller effect size as well as modest downregulation at 48hpi.

CCL20, CXCL16 and CCL3 were not significantly DE until 18-24hpi and were grouped into cluster 8 by maSigPro analysis. *CCL20* showing the strongest induction at log2FC 3.7 by 36hpi and *CXCL16* and *CCL3* reaching a maximum of 1.7 and 2 log2FC respectively at 36 and 24hpi.

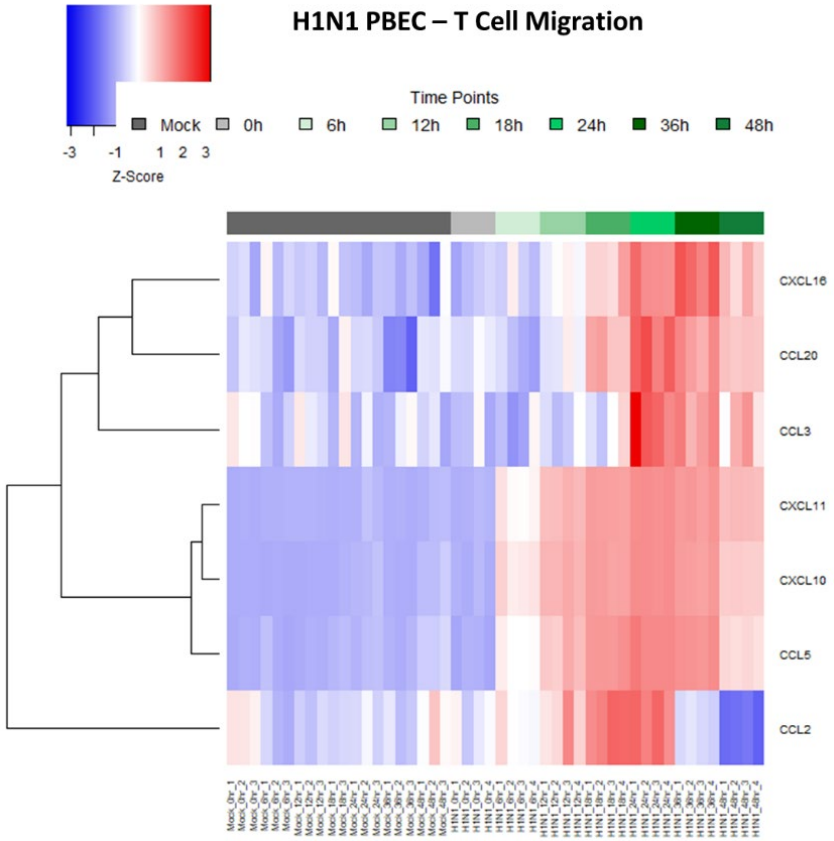


Figure 3-21 GSE47961 H1N1 caused strong upregulation of genes relating to PBEC - T cell Migration over 48h infection.

GSE47961 H1N1 expression data for 7 genes with relevance to PBEC modification of T cell migration. A heatmap was generated from neqc normalised intensity values for arrays from all H1N1 infected and mock-infected time points. Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Blue shading indicates that expression is up to 3 SD lower than the mean for a given gene and red shading indicates genes are expressed up to 3 SD above the mean.

Table 13 GSE47961 H1N1 Time-matched differential expression results for the PBEC - T cell migration gene set

	H1N1	6h	H1N1	12h	H1N1	18h	H1N1	24h	H1N1	36h	H1N1	48h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
CXCL10	7.162	4.64E-28	10.223	5.44E-37	11.089	8.49E-36	11.449	3.81E-40	11.408	8.18E-40	7.075	5.51E-28
CXCL11	5.346	1.18E-21	8.976	5.99E-34	10.514	7.21E-35	11.343	9.86E-40	10.891	3.97E-39	8.436	2.75E-32
CCL5	3.079	2.08E-17	5.368	4.31E-30	6.472	1.15E-31	7.033	5.67E-37	6.914	1.89E-36	3.089	9.26E-18
CCL20	0.101	9.61E-01	0.456	3.30E-01	1.694	1.40E-04	3.071	8.91E-13	3.669	1.01E-15	1.227	1.94E-03
CCL2	0.975	4.35E-04	1.436	3.79E-08	2.393	6.66E-14	2.228	3.11E-15	0.126	6.67E-01	-2.209	1.51E-14
CCL3	0.006	9.98E-01	-0.244	5.65E-01	0.105	8.51E-01	2.021	1.28E-09	1.186	1.42E-04	0.940	4.07E-03
CXCL16	-0.030	9.81E-01	0.405	9.52E-02	0.625	1.45E-02	1.448	4.32E-10	1.674	4.36E-12	1.034	5.05E-06

GSE47961 H1N1 Time matched DE analysis was performed on NEQC normalised data using Limma with results calculated for infected samples relative to mock-infected controls at every time point. Chemokine encoding genes were identified from genes significantly DE by H1N1 infection at 24hpi using ORA. Log2 fold change Log2FC values for these genes generated from time matched DE analysis of each time point relative to mock infected controls are shown here with FDR adjusted P values (FDR) coloured with a 3-point scale: Yellow > 0.05, Green: 0.05 – 1E-10, Purple: 1E-10 – 1E-40

A total of 20 genes were selected from the 85 genes mapped to “T cell activation” from the H1N1 gene list, no additional unique genes were present in the 19 which were identified from SARS-CoV1 upregulated genes. This included genes which encode MHC receptors also found in the term “Antigen processing and presentation” as well as surface expressed markers such as *CD200* which were also mapped to “T cell migration”. A heatmap of normalised expression values for these 20 genes throughout the H1N1 infection time series is included in **Figure 3-22**. Log2FC data from time matched DE analysis for genes encoding cytokines and surface molecules from this gene set are shown in **Table 14 and Table 15** respectively.

The greatest variation over time was seen for secreted cytokine encoding genes. *TNFSF13B* was upregulated from 6hpi reaching log2FC of 7.4 at 24hpi and remained around this level until a slight reduction to log2FC 4.9 by 48hpi. *IL-6* was also upregulated at 6hpi followed by an initial drop in upregulation at 12hpi and steady increase until 36h. This is similar to the DE of and *IL-1A* however, the 6 and 12hpi log2FC reading were not significant for this gene. *IL1B* and IL23a are only significantly DE at 24 and 36hpi, *IL7*, and *IL12a* are only minimally upregulated by H1N1 at all time-points. MHC encoding genes (*HLA-A*, *HLA-E*, *HLA-G* and *B2M*) were minimally upregulated between 18 and 48h with log2FC between 1 and 3. *CD274* was the only surface bound marker DE at 6hpi, and expression of this gene remains upregulated until the end of infection, with the strongest magnitude of DE for this category. Finally, *CD83* was also slightly upregulated between 6 and 12h but reaches more substantial levels at 18-36h time points while *CD55* is mildly upregulated at consistent levels throughout the infection.

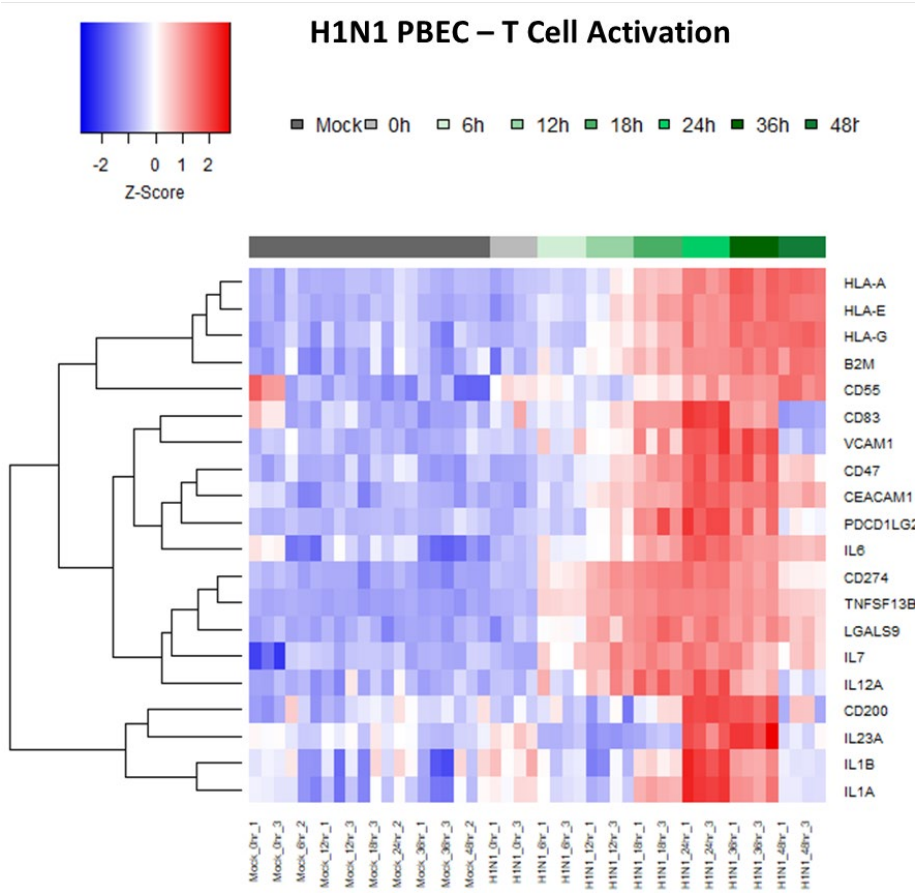


Figure 3-22 GSE47961 H1N1 caused strong upregulation of genes from the PBEC - T cell activation signature over 48h infection

GSE47961 H1N1 expression data for 20 genes with relevance to PBEC modification of T cell activation. A heatmap was generated from neqc normalised intensity values for arrays from all H1N1 infected and mock-infected time points. Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Blue shading indicates that expression is up to 3 SD lower than the mean for a given gene and red shading indicates genes are expressed up to 3 SD above the mean.

Table 14 GSE47961 H1N1 Time-matched differential expression results for the PBEC - T cell activation cytokine gene set

	H1N1	6h	H1N1	12h	H1N1	18h	H1N1	24h	H1N1	36h	H1N1	48h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
TNFSF13B	4.190	1.03E-21	6.107	9.40E-31	7.441	1.59E-32	7.269	2.50E-35	6.743	1.88E-33	4.945	1.17E-25
IL1A	1.022	1.38E-01	0.537	3.42E-01	2.361	1.21E-05	3.928	9.19E-14	3.834	2.54E-13	-0.086	9.19E-01
IL6	2.787	4.07E-12	0.909	1.25E-02	2.748	1.31E-10	3.845	7.16E-19	5.034	1.73E-24	3.971	6.02E-19
IL1B	0.623	7.74E-01	-0.310	7.72E-01	1.299	1.47E-01	3.287	4.12E-06	4.152	2.25E-08	-0.583	5.56E-01
LGALS9	1.131	1.95E-05	2.055	1.10E-13	2.531	2.65E-15	2.799	4.96E-20	2.471	1.59E-17	1.740	3.76E-11
IL23A	-0.373	4.51E-01	-1.324	8.74E-07	-0.651	2.91E-02	2.230	1.98E-14	3.747	1.14E-24	0.136	7.32E-01
IL7	0.757	1.33E-03	1.653	1.27E-12	1.725	1.60E-11	1.997	1.79E-16	1.716	8.97E-14	1.052	6.16E-07
IL12A	0.771	8.60E-03	1.277	5.65E-07	1.848	4.30E-10	1.809	5.64E-12	1.400	1.63E-08	0.245	4.27E-01

Genes encoding secreted cytokines DE by H1N1 infection at all time points. Values are arranged by descending log2 fold change (log2FC) at 24hpi. All fold change values are calculated for infected samples relative to time matched uninfected controls. Adjusted P values (FDR, were shaded using a 3-point scale, yellow = < 0.05

Table 15 GSE47961 H1N1 Time matched differential expression results for the PBEC - T cell activation surface bound gene set

	H1N1	6h	H1N1	12h	H1N1	18h	H1N1 24h		H1N1	36h	H1N1	48h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
CD274	2.315	1.19E-14	3.735	5.85E-25	4.834	4.28E-28	4.558	5.63E-30	4.440	2.92E-29	2.262	1.23E-14
CD83	0.892	2.16E-03	0.665	1.11E-02	2.602	6.01E-15	3.816	1.00E-25	2.754	1.03E-18	-0.259	4.11E-01
CEACAM1	0.844	1.25E-03	1.369	6.31E-09	2.748	3.47E-19	3.180	2.20E-25	3.958	9.89E-29	3.172	3.54E-22
VCAM1	0.708	4.19E-02	0.887	1.21E-03	1.736	3.42E-08	2.934	7.48E-19	2.829	4.34E-18	-0.187	6.17E-01
CD55	1.448	1.29E-02	0.975	1.43E-01	1.711	1.15E-02	2.054	2.21E-05	2.147	6.67E-05	2.053	9.28E-04
B2M	1.195	6.24E-02	1.698	2.75E-04	1.974	6.12E-05	2.017	3.24E-08	2.205	1.13E-12	2.066	8.88E-09
HLA-E	0.373	2.10E-01	0.891	4.76E-04	1.582	4.23E-09	1.807	1.27E-12	2.301	5.20E-16	2.157	8.73E-16
HLA-A	0.262	6.20E-01	0.641	2.74E-02	1.457	3.85E-06	1.763	1.67E-09	2.397	1.27E-13	2.329	3.47E-13
CD47	0.285	4.76E-01	0.684	1.38E-03	1.329	6.70E-07	1.712	1.36E-13	1.673	1.31E-07	0.745	1.05E-02
PDCD1LG2	0.110	7.49E-01	0.591	2.62E-06	1.366	1.49E-16	1.709	4.37E-24	1.333	8.16E-19	0.230	7.57E-02
HLA-G	0.408	2.74E-01	0.567	2.18E-01	1.155	2.50E-05	1.483	3.18E-11	1.307	8.80E-08	0.883	1.84E-03
CD200	0.024	9.87E-01	-0.041	9.25E-01	0.181	6.61E-01	1.162	8.52E-06	1.246	2.50E-06	0.053	9.10E-01

Genes encoding surface bond receptors relating to T cell activation were DE by H1N1 infection at all time points. Values are arranged by descending log2 fold change (log2FC) at 24hpi. All fold change values are calculated for infected samples relative to time matched uninfected controls.
Adjusted P values (FDR), were shaded using a 3 point scale: Yellow > 0.05, Green: 0.05 – 1E-10, Purple: 1E-10 – 1E-40

Although SARS-CoV1 did not result in significant enrichment of T cell terms, there were genes with roles in these pathways DE by this virus. None were unique to SARS-CoV1, however, SARS-CoV1 infection did have an impact on several genes from the H1N1 PBEC – T cell signature. Despite not being mapped to "T cell migration" during the clusterProfiler ORA, both *CXCL10* and *CXCL11* were significantly DE by SARS-CoV1, again, being among the most highly upregulated by this virus reaching log2FC 5.6 and 5.1 respectively at 72hpi. As no other genes from "T cell migration" were DE by SARS-CoV1, *CXCL10* and *CXCL11* have been included with "T cell activation" for presentation here as a heatmap **Figure 3-23,** and log2FC values and adjusted P values for time matched DE analysis **Table 16.** *CXCL10* and *CXCL11* were expressed with the cluster 4 maSigPro pattern as well as MHC encoding genes (*HLA-E, HLA-A* and *B2M*), *LGALS9* and *TNFSF13B*. The HLA genes however were not DE with any significance. HLA-A was the only gene from this family to exceed log2FC 1 and only did so at 84hpi. Similarly, *LGALS9* is DE with log2FC 1.2 at 72 and 84hpi only while *TNFSF13B* is upregulated at 48h and remains so throughout infection but with DE reaching a maximum of 3.2 log2FC at 84hpi.

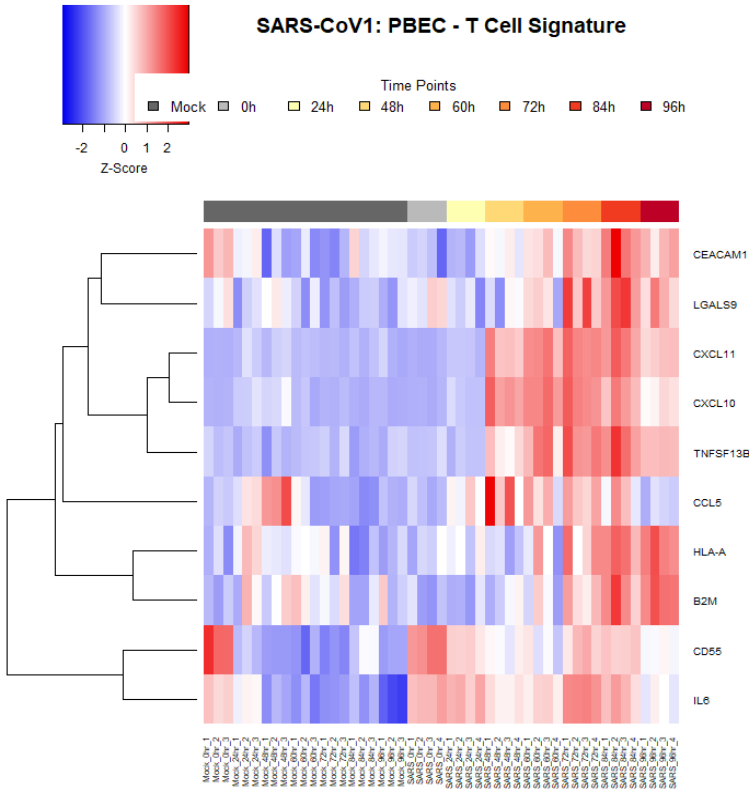


Figure 3-23 GSE47961 SARS-CoV1 infection had a varied impact on expression of genes from the PBEC - T cell gene signature over 72h infection

GSE47961 SARS-CoV1 expression data for all genes with relevance to PBEC modification of T cell activation with CXCL10 and CXCL11 included from T cell migration. A heatmap was generated from neqc normalised intensity values for arrays from all infected and mock-infected time points. Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Blue shading indicates that expression is up to 3 SD lower than the mean for a given gene and red shading indicates genes are expressed up to 3 SD above the mean.

Table 16 GSE47961 SARS-CoV1 Differential expression results for the PBEC - T Cell recruitment/activation gene set

	SARS1	24h	SARS1	48h	SARS1	60h	SARS1	72h	SARS1	84h	SARS1	96h
SYMBOL	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**	**Log2FC**	**FDR**
CXCL10	-0.273	9.49E-01	3.995	1.45E-14	5.395	9.61E-21	5.683	1.85E-22	4.735	1.34E-18	2.865	1.80E-09
CXCL11	-0.108	9.73E-01	3.336	1.48E-11	4.326	1.00E-15	5.156	1.99E-20	4.861	2.51E-19	3.188	2.66E-11
IL6	1.507	2.20E-03	2.363	7.40E-09	2.799	1.54E-11	4.574	2.45E-21	3.087	1.68E-13	3.422	3.41E-15
TNFSF13B	0.065	9.97E-01	1.652	9.10E-06	2.252	1.33E-09	2.688	1.08E-12	3.227	5.17E-16	1.823	6.81E-07
CD55	1.465	1.68E-02	1.425	3.90E-03	1.264	2.12E-01	1.944	1.83E-03	1.078	1.44E-01	1.064	2.16E-01
LGALS9	0.045	9.97E-01	-0.213	6.03E-01	0.373	5.49E-01	1.192	1.20E-05	1.156	2.37E-05	0.803	1.38E-02
HLA-A	-0.218	8.50E-01	-0.135	7.30E-01	-0.035	9.99E-01	0.495	2.26E-01	1.043	8.24E-04	0.913	9.47E-03

Log2 fold change (log2FC) values calculated using pairwise differential expression analysis of SARS-CoV1 infected PBEC ALI relative to mock infected control are shown for genes mapped to the GO:BP term "T cell Activation". Adjusted P values (FDR) are corrected for false discovery of multiple hypothesis testing using eBayes (limma). Values are coloured by significance in a 3-point scale with yellow indicating AdjP > 0.05.

Most cytokine encoding genes from the H1N1 T cell activation signature were not significantly DE by SARS-CoV1. Only *IL6* was DE by SARS-CoV1 in the time-matched analysis. This gene was expressed with the cluster 6 expression pattern and was upregulated log2FC 1.5 at the earliest 24h time point

and reached a maximum upregulation of log2FC 4.6 at 72hpi. The GSE47961 PBEC – T cell gene signature was comprised of 27 genes, 5 of which were not identified from the GSE32138 analysis: *TNFSF13B*, *IL7*, *CD55*, *PDCD1LG2* and *CXCL16.* All genes were identified from the H1N1 analysis; only 7 genes were also DE by SARS-CoV1 at the chosen thresholds (log1FC>1, FDR < 0.05) with no additional genes specific to SARS-CoV1 infection.

3.2.3 Combined PBEC – T cell gene signature

Combining unique genes from both data sets produced a final PBEC - T cell gene signature comprised of 41 genes with mixed expression in response to each virus. The GSE47961 time-points which produced the most DE were chosen for comparison with the GSE32138 arrays. For H1N1 this was 24hpi, for SARS-CoV1 this was 72 and 84hpi. H1N1 48h arrays were also included to provide time matched comparison with RSV. SARS-CoV1 24 and 48h arrays were not included due to the lack of DE at these time points. Normalised expression data for these samples were combined and plotted as a single heatmap **(Figure 3-24).** In this plot, columns were clustered using wards algorithm to determine the natural grouping of the data. Most arrays clustered by treatment group with some overlap between the two SARS-CoV1 time points. As with previous results, data from pairwise DE analysis for these genes are included as tables for each sub category: PBEC- T cell recruitment **Table 17,** PBEC – T cell activation cytokines **Table 18** and PBEC – T cell activation surface molecules **Table 19.**

Clustering of the data by sample identified strong similarity between the H3N2 and H1N1 24h IAV infections which were grouped together and away from the other arrays. Both infections resulted in near identical upregulation of chemokine genes *CXCL10*, *CXCL11*, *CCL5*, and *CXCL16* at 24hpi, while *CCL20* and *CCL2* were more abuntantly upregulated by H3N2 . Genes encoding several surface bound T cell activating proteins were also comparably expressed by both strains of IAV at 24h, particularly the HLA genes which were also similarly weakly upregulated by RSV. H3N2 caused greater

upregulation of *CD274*, *CEACAM1*, *CD200*, *CD70*, *PDCD1LG2* and to a lesser extent *CD276* than H1N1. Secreteted cytokine DE was even more varied with H3N2 resulting in particularly high upregulation of *IL6*, *IL1A*, *IL1B*, *TNF* and *IL23A* compared to H1N1.RSV caused DE of only , *IL6*, *TNFSS13B*, *LGALS9* and *IL15* from the cytokine subset of this signature. These genes were also DE by SARS-CoV1 at 72hpi and, with the exceptioon of *IL6*, to similarly low levels by each virus. The response to RSV and H1N1 at 48hpi, were also fairly similar in that by this time point, H1N1 infection has a comparably weak impact on most genes discussed.

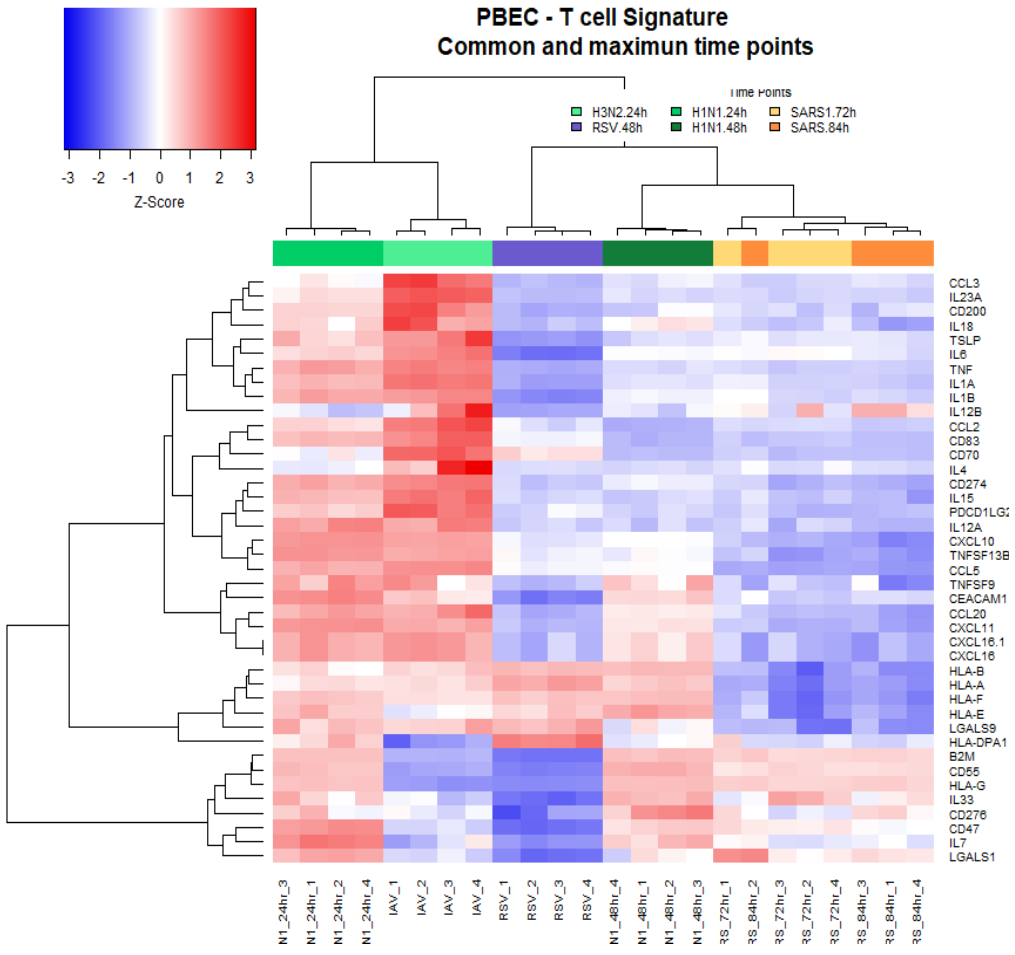


Figure 3-24 PBEC - T cell gene signature is strongly expressed in response to IAV but less so in response to RSV or SARS-CoV1

Expression values for averaged replicate arrays from time points with greatest expression or that are shared between datasets. Genes (rows) were scaled as Z-scores showing standard deviation from the mean expression of each gene: For this plot, Blue shading indicates that expression is up to 3 SD lower than the mean for a given gene and red shading indicates genes are expressed up to 3 SD above the mean. Coloured bars indicate group membership as shown in the key.

Table 17 PBEC-T cell Recruitment gene expression in response to all four viruses

	RSV	48h	H3N2	24h	H1N1	24h	H1N1	48h	SARS1	72h	SARS1	84h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
CXCL11	6.800	3.44E-12	12.087	1.95E-16	11.343	9.86E-40	8.436	2.75E-32	5.156	1.99E-20	4.861	2.51E-19
CXCL10	7.601	1.62E-14	11.973	3.35E-17	11.449	3.81E-40	7.075	5.51E-28	5.683	1.85E-22	4.735	1.34E-18
CCL5	4.009	2.50E-11	8.905	8.13E-16	7.033	5.67E-37	3.089	9.26E-18	1.086	1.87E-03	0.968	8.12E-03
CCL3	0.054	9.39E-01	8.227	1.47E-11	2.021	1.28E-09	0.940	4.07E-03	-0.039	9.97E-01	0.525	3.98E-01
CCL20	2.434	1.44E-03	6.527	5.55E-09	3.071	8.91E-13	1.227	1.94E-03	1.098	5.00E-02	0.833	2.09E-01
CCL2	-0.060	9.52E-01	6.488	1.14E-08	2.228	3.11E-15	-2.209	1.51E-14	0.276	7.90E-01	0.024	9.80E-01
CXCL16	0.282	1.07E-01	1.947	9.11E-10	1.448	4.32E-10	1.034	5.05E-06	0.109	9.68E-01	0.167	7.69E-01

Differential expression results for chemokine encoding genes from the PBEC – T cell signature are shown for all four viruses. The GSE47961 timepoints which overlap with those used for the GSE32138 data as well as the timepoints with the greatest impact on gene expression were chosen for comparison. Log2 fold change (Log2FC) and FDR adjusted P values are arranged by descending log2FC for H3N2 24h. FDR significance values are coloured with a 3 point scale: yellow = > 0.05.

Table 18 PBEC - T cell activation, cytokine gene expression in response to all four viruses

	RSV	48h	H3N2	24h	H1N1	24h	H1N1	48h	SARS1	72h	SARS1	84h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
IL6	1.213	1.42E-03	10.482	1.02E-14	3.845	7.16E-19	3.971	6.02E-19	4.574	2.45E-21	3.087	1.68E-13
IL1A	0.556	1.08E-02	8.875	2.89E-16	3.928	9.19E-14	-0.086	9.19E-01	0.612	7.13E-01	0.695	4.63E-01
IL1B	0.146	7.01E-01	8.584	4.50E-14	3.287	4.12E-06	-0.583	5.56E-01	0.455	9.48E-01	0.720	6.68E-01
TNF	0.361	8.24E-01	7.109	3.03E-14	2.461	2.10E-17	0.149	6.81E-01	0.449	4.65E-01	0.448	4.74E-01
IL23A	0.470	3.76E-02	7.090	7.08E-15	2.230	1.98E-14	0.136	7.32E-01	0.222	9.00E-01	0.423	3.92E-01
TNFSF13B	4.880	4.26E-13	6.843	2.73E-15	7.269	2.50E-35	4.945	1.17E-25	2.688	1.08E-12	3.227	8.69E-01
TSLP	0.062	9.03E-01	2.734	1.10E-07	1.139	3.96E-05	-0.040	9.39E-01	-0.153	9.61E-01	0.141	5.17E-16
IL33	0.255	6.49E-01	2.633	3.31E-06	1.941	6.42E-04	2.391	6.76E-05	0.728	7.47E-01	0.383	8.12E-01
LGALS9	2.377	6.38E-04	2.534	5.61E-05	2.799	4.96E-20	1.740	3.76E-11	1.192	1.20E-05	1.156	2.37E-05
IL12A	0.291	1.34E-01	2.011	2.09E-09	1.809	5.64E-12	0.245	4.27E-01	0.183	9.28E-01	0.255	6.40E-01
IL15	1.313	1.42E-07	1.737	5.33E-10	1.420	4.48E-14	0.486	4.22E-03	0.146	8.96E-01	0.564	7.82E-03
LGALS1	0.200	6.67E-01	1.604	7.58E-05	-0.272	6.76E-01	-2.041	6.45E-04	-0.356	9.38E-01	-0.037	9.77E-01
IL12B	-0.096	8.55E-01	1.529	1.44E-04	0.114	7.59E-01	0.336	3.62E-01	0.242	9.14E-01	1.049	5.08E-03
IL4	0.001	9.97E-01	1.088	9.88E-05	0.115	2.88E-01	0.003	9.90E-01	0.055	9.61E-01	0.051	8.65E-01
IL18	-0.022	9.48E-01	1.065	3.24E-05	0.264	4.70E-02	0.308	2.96E-02	0.015	9.98E-01	0.005	9.93E-01
IL7	0.794	3.37E-02	0.988	2.89E-03	1.997	1.79E-16	1.052	6.16E-07	0.092	9.74E-01	0.371	3.33E-01

Differential expression results for secreted cytokine encoding genes from the PBEC – T cell signature for all four viruses. The GSE47961 time points with the greatest Impact on gene expression or which overlap with the GSE32138 data, were chosen for comparison. Log2 fold change (log2FC) and FDR adjusted P values (FDR) are Arranged by log2FC for the H3N2 24h data. Significance values are coloured with a 3 point scale: yellow = > 0.05.

Table 19 PBEC - T cell activation, surface bound, gene expression in response to all four viruses

	RSV	24h	H3N2	24h	H1N1	24h	H1N1	48h	SARS1	72h	SARS1	84h
SYMBOL	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR	Log2FC	FDR
CD274	1.904	3.94E-05	6.221	5.03E-12	4.558	5.63E-30	2.262	1.23E-14	1.093	1.68E-04	1.175	4.09E-05
CEACAM1	2.870	2.61E-08	5.742	3.46E-13	3.180	2.20E-25	3.172	3.54E-22	0.842	1.43E-04	0.987	1.54E-04
CD83	0.988	1.99E-02	5.143	2.59E-10	3.816	1.00E-25	-0.259	4.11E-01	0.549	2.16E-01	0.460	3.24E-01
CD200	0.131	7.52E-01	3.397	3.13E-09	1.162	8.52E-06	0.053	9.10E-01	0.184	9.38E-01	0.086	9.22E-01
CD70	1.696	2.42E-06	3.223	1.22E-10	0.863	2.74E-04	-0.627	1.34E-02	0.043	9.94E-01	0.171	7.94E-01
PDCD1LG2	0.638	6.30E-03	3.107	5.88E-11	1.709	4.37E-24	0.230	7.57E-02	0.052	9.74E-01	0.196	4.00E-01
B2M	1.259	4.53E-05	2.908	3.88E-04	2.017	3.24E-08	2.066	8.88E-09	0.558	3.61E-01	0.827	9.71E-04
CD55	1.529	5.59E-01	2.873	8.98E-05	2.054	2.21E-05	2.053	9.28E-04	1.944	1.83E-03	1.078	1.44E-01
HLA-G	2.427	4.21E-03	2.571	1.45E-03	1.483	3.18E-11	0.883	1.84E-03	0.397	3.14E-01	0.342	5.72E-01
CD47	1.415	8.94E-03	2.045	8.10E-08	1.712	1.36E-13	0.745	1.05E-02	0.361	4.86E-01	0.379	3.51E-01
HLA-F	1.909	4.61E-09	1.843	1.79E-08	2.399	2.52E-21	2.636	8.41E-24	0.697	4.07E-04	1.217	1.80E-10
HLA-E	1.782	4.79E-05	1.444	1.50E-05	1.807	1.27E-12	2.157	8.73E-16	0.484	1.50E-01	0.910	6.83E-04
HLA-A	1.883	7.39E-09	1.429	2.08E-08	1.763	1.67E-09	2.329	3.47E-13	0.495	2.26E-01	1.043	8.24E-04
HLA-B	1.251	1.24E-07	1.179	2.57E-07	1.678	7.51E-11	2.208	2.25E-15	0.726	5.23E-03	1.159	2.04E-06
CD276	0.349	9.43E-01	1.141	4.02E-05	0.366	5.92E-02	0.495	1.78E-02	0.045	9.90E-01	0.347	3.81E-01
HLA-DPA1	1.236	1.18E-02	-0.678	4.81E-02	0.509	3.22E-02	-0.087	8.16E-01	0.029	9.96E-01	0.081	9.05E-01

Differential expression results for surface expressed molecule encoding genes from the PBEC – T cell signature for all four viruses. The GSE47961 time points with the Greatest Impact on gene expression, or which overlap with the GSE32138 data, were chosen for comparison. Log2 fold change (log2FC) and FDR adjusted P values (FDR) Are arranged by descending log2FC for the H3N3 24h data. Significance values are coloured with a 3-point scale: yellow = > 0.05.

3.3 Discussion

This chapter explored transcriptomic changes induced by four respiratory viruses, H3N2, RSV, H1N1 and SARS-CoV1, in ALI models of airway infection. There were considerable differences in the number of genes DE by each virus as well as both similarities and differences in the biological processes influenced by each infection. Genome wide DE analysis identified significant enrichment of immune system related pathways in genes upregulated by all four viruses. Genes relating to innate responses such IFN stimulated genes, were significantly DE by each virus to broadly similar extents. These antiviral gene pathways dominated the response generated by SARS-CoV1 and RSV and made up the bulk of the relatively small number of genes DE by these viruses. Both H1N1 and H3N2 produced a considerably stronger and more varied transcriptomic response with many more cytokine encoding genes upregulated by each virus than was seen for RSV or SARS-CoV1. IFN inducible gene expression was broadly similar but the overall transcriptomic response to H1N1 and H3N2 was more varied and had a larger effect size than was seen in response to RSV or SARS-CoV1 infection.

Exploring the temporal changes in gene expression in response to H1N1 and SARS-CoV1 identified that for H1N1, antiviral gene expression begins immediately with significant DE of genes with roles in antiviral immunity occurring as early as 6h post infection. In contrast, antiviral gene expression begins to occur 48h after SARS-CoV1 infection and required 3 days for many genes to be upregulated. GO:BP terms relating to the inflammatory response and cytokine expression were significantly enriched in genes DE by H1N1 at slightly later time points, 12-18hpi, and were maintained at peak levels of upregulation between 24 and 36h, while antiviral gene expression declines from a 24h peak until the 48h end of infection. This clear distinction between immediate and later immune responses was not as evident for SARS-CoV1 infection due to the small number of genes DE by this virus. The small number of genes downregulated by SARS-CoV1 also complicated the comparison with H1N1 as a temporal profile of gene inhibition could not be identified for SARS-

CoV1. Looking at only 48hpi downregulation however, there were similarities between both GSE47961 data sets and the H3N2 data. SARS-CoV1 caused downregulation of genes for extracellular matrix proteins and both strains of IAV resulted in downregulation of similar genes with additional enrichment of cilia function and formation related GO:BP terms. For H1N1, these pathways remained consistently downregulated from 24hpi until the 48h end pint. This may point towards loss of apical cilia during infection that was not seen in response to RSV and was less pronounced in response to SARS-CoV1. A similar study utilising ALI differentiated primary airway cells reported that H3N2 resulted in large-scale ciliated cell destruction that did not occur in response to other viruses (5 subtypes of Rhinovirus (RV), RSV-B and HCoV-OC43) (118). Biological process terms such as "regulation of metabolic process" were also downregulated by both viruses, however in response to H1N1, expression levels for genes from these terms return to baseline by 48hpi, suggesting a recovery of cellular metabolic functions.

With only a single time point provided for arrays in the GSE32138 data set, the temporal dynamics of H3N2 and RSV could not be explored here. The data from this 24h H3N2 infection was similar to the H1N1 arrays from the same timepoint despite a 50% lower initial viral load with H3N2 infected at an MOI of 1 and H1N1 an MOI of 2. While the variations in NA and HA which define the IAV subtypes can impact infection, the replication kinetics of different isolates of H1N1 and H3N2 virus have been shown to be broadly similar during *in vitro* infection of differentiated human bronchial epithelial cells (119). Using an *in vivo* infection model of the ferret lung (120), different isolates of H3N2 produced comparable infection of nasal turbinates but only the H3N2 (A/Udorn/307/1972) virus, also used in the GSE32138 dataset, was able to efficiently replicate in the lower airways. As the models used here utilised only upper airway cultures, it is possible that affinity for infection was comparable between these viruses.

Both strains of IAV also resulted in stronger upregulation of type III IFNS than Type I, supporting research in which type III IFNs are produced at higher levels and during longer times in the lung than

type I IFNs during influenza virus infection (62). GSE32138 RSV also preferentially induced type III IFN expression but to a much lesser extent than the IAV strains. This may reflect the biological reality, for example nasal washes from RSV infected infants contain hardly any IFN-α and IFN –β in comparison to samples from other viral infections, including IAV (121). However, the 48h incubation period may also have been suboptimal as RSV is thought to replicate relatively slowly, reaching peak concentrations 4 days into culture, possibly due to slower spreading between cells (122). It is also possible that the low IFN expression seen response to infection are an indication inhibited IRF signalling due to virial evasion strategies. For example, the non-structural proteins of IAV (IAV-NS1) prevent the activation of several ISGs and the nuclear translocation of NF-κB and IRF3, preventing IFN production, while RSV-NS1 and NS2 suppress both the production and signalling of type I and III IFNs (123,124). In experiments with A549, 293T, and HEp-2 human epithelial cells, *in vitro* infection with RSV-A2 or various clinical isolates, failed to induce IFN-a production beyond the levels seen in mock infected controls. The authors also demonstrated that RSV infection supressed ongoing type I IFN signalling when administered to cells following stimulation with a TLR9 agonist.

Relatively weak innate immune induction has also been seen for several strains of CoV, suggesting an ability to remain undetected by innate immune sensors. One study using an ALI infection model demonstrated that cultures infected with SARS-CoV-2 exhibited the lowest expression of 3 ISGs when compared with those infected by IAV or RSV (125). Additionally, the Nsp3 protein from pandemic CoVs inhibits IFNB production following TLR7 activation and Nsp1 proteins can severely inhibit host protein translation, including IFNs, via interaction with the 40S ribosomal subunit (121).

The PBEC – T cell gene signature developed from this project aimed to compare biologically relevant gene expression across viral groups. Many genes within T cell related biological process terms were not relevant to PBEC biology instead describing pathways such as TCR expression. To develop the PBEC – T cell signature, genes within these terms which were DE by each virus were selected based on whether their products are expressed by airway epithelial cells and able to interact with T cells.

Chemokine encoding gene expression has perhaps the most direct impact on immune cell regulation by airway epithelial cells. However, for most chemokines there are multiple possible target receptors expressed by a range of leukocytes, complicating the interpretation of these findings.

All viruses produced strong induction of CXCL10 and CXCL11, chemokines which bind to and activate the receptor CXCR3 expressed on the surface of monocytes, T cells and NK cells. CXCL11 is considered the most potent activator of CXCR3 and induces trans-epithelial migration of T cells. Both chemokines however are associated with Th1 primed T cell responses. CCL5, CCL4 and CCL3 are all ligands for CCR5, while CCL20 interacts with CCR6. CCL3 and CCL20 were expressed to different extents by the viruses discussed in this project, showing greatest upregulation by H3N2. CCL20 is the sole ligand for the chemokine receptor CCR6, selectively expressed on immature DCs as well as B cells and subsets of CD4+ T cells including effector memory, Th17 and Treg cells (126). A study using CCR6-/- mice found defective DC recruitment following infection with RSV as well as reduced lung pathology (127). The CCR6 deficient mice were also better able to clear infection and exhibited a Th1 dominant immune response in contrast to the mixed Th1/Th2 T cell response by WT mice. The relatively week expression of this cytokine in response to SARS-CoV2 infection may point to excessive Th1 driven responses. CCL3 is chemotactic for both T cells and NK cells. This chemokine has been linked to RSV induced pathology in a murine study using CCL3 knockout mice (128). The authors found that CCL3 regulates T cell populations with CCL3 inhibition resulting in higher numbers of RSV-specific pro-inflammatory T cells being recruited to the lung.

The expression of cytokine encoding genes from the PBEC-T cell signature also differed quite considerably between viruses, especially concerning *IL6*, *IL23A* and *IL1B*. In addition to showing stronger induction by the IAV strains, these genes were also differently upregulated by H1N1 and H3N2, with the latter showing stronger induction. *IL6* was upregulated with a 10.5 log2 fold change by H3N2 but only log2FC 3.8 by H1N1 at the same 24h time point. Neither RSV nor SARS-CoV1 infection resulted in differential expression of *IL1B* or *IL23A* and RSV also had no impact on *IL6*.

SARS-CoV1 actually resulted in IL6 upregulation at 72h which was stronger than that seen for H1N1 at 24h. No other gene from the T cell activation signature was upregulated by SARS-CoV1 to his extent. Interestingly, high levels of IL-6 have been linked to patient decline in COVID-19 with anti-IL6 drug Tocilizumab licenced for treatment (129).

These cytokines have numerous roles in the immune response but may be key drivers of Th2 type responses. *IL23A* encodes a subunit of the heterodimeric cytokine IL-23 which is associated with Th17 differentiation and maintenance, most efficiently in conjunction with IL-1β. IL-6, has also been linked to Th17 development through induction of IL-21. When considered alongside the differences in *CCL20*, this pattern suggests that H3N2 in particular might induce stronger Th17 differentiation and recruitment than the other viruses. Interestingly, cytokines (IL-6 and CXCL8) released from SARS-CoV1 infected ALI differentiated Calu-3 cells have been shown impair DC and macrophage activation and priming of naïve CD4+ T cells (130). As in this project, the authors also reported a delay in cytokine production with IL6 not becoming significantly distinct form controls until 4 days post infection, despite infectious SARS-CoV1 progeny reaching a maximum after 2 days.

There was more similarity between viruses for many of the surface bound molecules identified in this analysis. Antigen presentation via MHC receptors is a key mechanism by which airway epithelial cells might influence T cells. In this work however, no viruses caused substantial DE of HLA genes and there were no significant differences between the four viruses. CD83 was most induced by H3N2 infection and not upregulated by RSV or strongly induced by SARS-CoV1. This gene encodes a surface receptor most commonly associated with dendritic cells, however a soluble form of the encoded protein has recently been shown to suppress antigen-specific Th2 response in the nasal mucosa (131). *CEACAM1* and *CD274* point to regulation of CD8+ T cell activation and inhibition and were also most prominently upregulated by H3N2 with minimal DE by RSV or SARS-CoV1. The product of *CD274* (PDL-1) is a central component of CTL inhibitory signalling via PD-1. Interactions between PDL-1 and PD-1 may contribute to the functional inactivation of virus-specific CD8+ T cells during

chronic viral infection (132). CD274 is upregulated immediately by H1N1 but with greatest magnitude by H3N2 and is only minimally DE by both RSV and SARS-CoV1 (at 72h). *CEACAM1* encodes a transmembrane protein expressed by epithelial cells at mucosal sites as well as some immune cells (126). It functions as an attachment molecule for bacteria and the two identified isoforms of the transmembrane *CEACAM1* domain have been described to either activate T cells and induce Treg formation or inhibit activated T-cell function. *CEACAM1* may be essential for CD8+ T cell activation as the absence of *CEACAM1* on virus-specific CD8+ T cells limits the antiviral CD8+ T cell response in both in vivo murine lymphocytic choriomeningitis virus (LCMV) infection and human Peripheral Blood Mononuclear Cells (PBMC) (133). *CEACAM1* can be activated by contact with other *CEACAM1* molecules so a role for direct epithelial – T cell activation is an interesting possibility.

The similarities seen for both strains of IAV despite the experimental work being performed by different researchers, using different PBEC and different strains of IAV, supports the validity of comparing DE across these data sets. The use of the same microarray platform in each data set improved the direct comparability however, this work could benefit from using additional methods of meta-analysis in which the data are combined prior to shared normalisation and DE analysis. Despite this, the work presented here provides an interesting comparison of these four common respiratory viruses.

Chapter 4 In vitro infection model development

4.1 Introduction

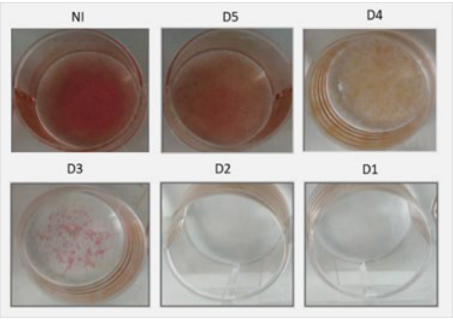
Most RSV infection models in the literature utilise immortalised cell lines such as VERO-E6 or Hep-2 cells in 2D cultures. However, 2D models cannot represent the airway architecture beyond the responses of a single cell type and frequently lack comparable viral entry receptor expression. Immortalised cell lines are also distinct from primary cells and may show greater susceptibility to viral infection which does not reflect biological reality. Well differentiated airway epithelial cell models re-create several aspects of the in vivo airway including pseudostratified epithelium, cilia beating and mucus production. Human primary bronchial epithelial cells (PBEC) are not always available and are prone to donor related variability. A secondary goal of this work was to establish an RSV ALI infection model using the recently developed immortalized human airway basal cell line BCi-NS1.1. These cells retain characteristics of the originating primary cells over long-term culture. Importantly, they show multipotent differentiation capacity producing ciliated epithelial cells as well as goblet cells in ALI culture. If successful, this model would combine the biological validity of a polarised epithelial cell model with the consistency of a cell line.

The viability of laboratory strain RSV-A2 was confirmed using the known RSV permissive cell line VeroE6. This was followed by imaging of RSV infection in monolayer cultures of BCi cells prior to ALI culture infections. In the 3D models, cells from both the large and small airways were used with a view to identifying differences in response to RSV. Successful modelling of both airway compartments using this model would offer important insight into the pathology of RSV. No studies have been found that demonstrate RSV infection in ALI cultures of the large airway BCi cell line or the small airway BCi. For this project, RSV infection was carried out with a 72h maximum post infection incubation period. The protocol used was based on previous work using this BCi model for IAV infection as well as studies showing that RSV has a slower rate of growth than IAV both clinically

and in vitro (122). This chapter describes the steps taken to fulfil these goals and complete aims 1-2 of this project.

4.2 RSV visualisation in monolayer cell infections

Initial work with monolayer cultures of VERO cells was able to show viability of the new RSV-A2 strain. The first plaque assay **(Figure 4-1**) showed complete cell death in response to the first and second serial dilutions of RSV containing $3.5x10^5$ and $1.75x10^5$ PFU of virus. The amount of cell lysis reduces with decreasing viral concentration however, there was no point at which defined plaques, regions of cell death constrained to discreet foci, were visible. RSV produces relatively small plaques in contrast to other viruses. This may be a factor here however, the lowest concentrations of virus D4 ($4.37x10^4$ PFU) and D5 ($2.18x10^4$ PFU) offer the best possibility of observing plaque formation or the presence of multinucleated cells (syncytia) which would be a more definitive indication of cytopathic effect (CPE)



	RSV (µl)	Diluent (µl)	PFU (50µl)
Dilution 1	50 Stock	450	$3.5x10^5$
Dilution 2	250 D1	250	$1.75x10^5$
Dilution 3	250 D2	250	$8.75x10^4$
Dilution 4	250 D3	250	$4.37x10^4$
Dilution 5	250 D4	250	$2.18x10^4$

.

Figure 4-1 VERO cell monolayers were highly susceptible to cell lysis following infection with serial dilutions of RSV-A2.

VERO E6 cells were grown to confluence on 12 well plates before infection with 2-fold serial dilutions of RSV. An overlay consisting of 0.3% agarose was added to the cells after a 2h infection period to prevent viral movement through the supernatant. Restricting spread to cell-to-cell contact encourages plaque formation as infected cells lyse. RSV produces relatively small plaques in VERO cells however, it is unlikely that plaque formation happened here. There was considerable cell death at the highest concentration D1 and D2 dilutions as well as substantial cell death in D3 and D4 infected cells. Calculated PFU for each 2-fold serial dilution of RSV-A2 are included in the accompanying table.

Broad cell lysis could be explained by inadequate restriction of viral movement using the 0.3% agarose overlay. The experiment was therefore repeated using only the D4 and D5 serial dilutions with overlays consisting of 0.5% agarose or Avicel (microcrystalline cellulose) to increase the likelihood of plaque formation. Widespread lysis was not observed in this experiment and small defined plaques were visible in all infected wells, most clearly in wells with Avicel overlay and the D5 RSV-A2 dilution **(Figure 4-2 Arrow**). Clumps of apoptotic cells and multi nucleated giant cells (syncytia) were visible throughout the infected VERO cell monolayer (**Figure 4-2 B Circle**).

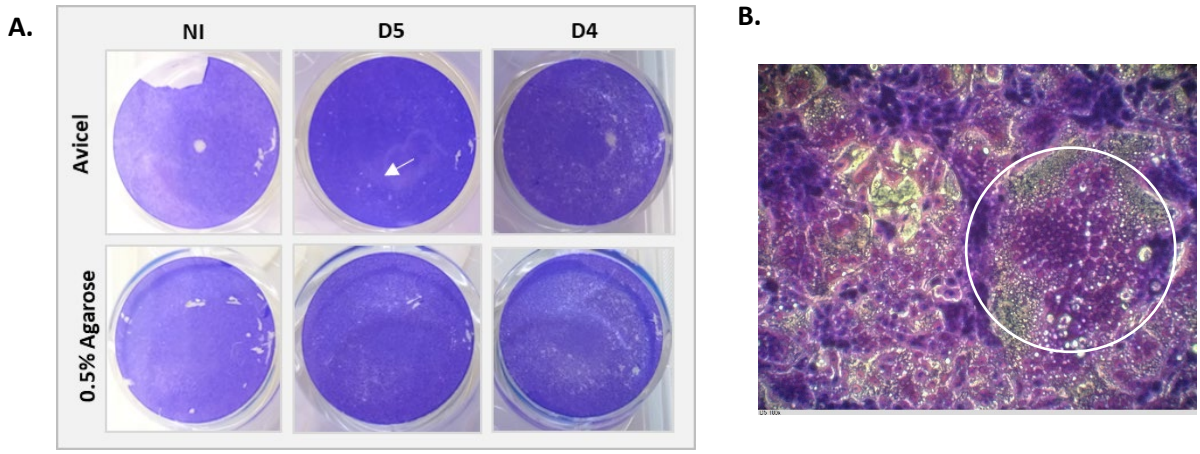


Figure 4-2. Low concentrations of RSV-A2 resulted in defined plaque formation in VERO cell monolayers using Avicel overlay medium.

In image A, uninfected cells (NI) show some damage to the cell layer post staining but no plaques visually or microscopically. D5 infection produced clear plaques (arrow). More plaques are present in the D4 infection but less clearly defined. Image B. Microscopic imaging of plaques formed in the D5 Avicel well. 1000x Image is representative of duplicate wells. A tightly packed ball of cells (syncytium) fused together by RSV-F with evidence of apoptosis is circled. These cells lyse to produce the small regions of absent staining referred to as plaques.

Based on this plaque assay, RSV-A2 was used at the 3rd, 4^{th,} and 5th serial dilutions, 8.75x10^4, 4.37x10^4 and 2.18x10^4 PFU respectively, per well for BCi monolayer infection. Representative images of RSV-G immuno-fluorescent imaging of this infection at 100x and 200x magnification are shown alongside an uninfected sample at 100x magnification in **Figure 4-3**.

RSV-G was primarily detected in groups of cells along the edges of gaps in the cell monolayer. There was no visible difference in the number of cells staining positive for RSV-G with each concentration

of virus. At 200x magnification some groups of cells are surrounded by the same RSV-G covered cell membrane.

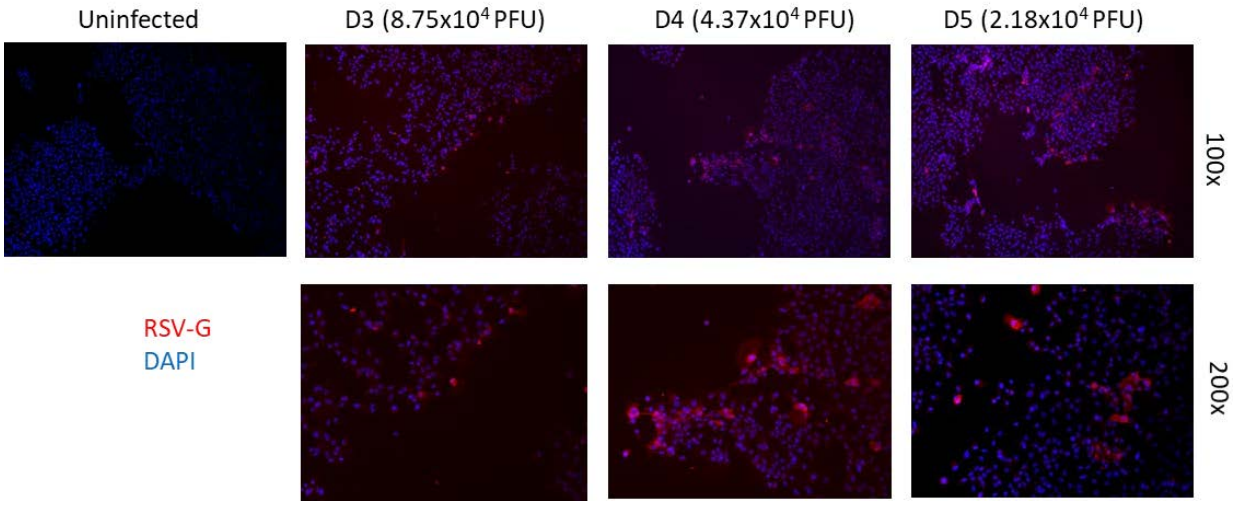


Figure 4-3 Immuno-fluorescent staining of RSV-G protein in BCi-NS1.1 2D cultures BCi 2D culture infected with low concentrations of RSV.

BCi monolayers cultured in 4 well chamber slides were infected with low concentrations of RSV before fixation and immuno-fluorescent imaging of RSV-G protein using Goat anti-Rabbit AF-568 conjugated secondary antibodies. Cell Nuclei are stained with DAPI (BLUE). No red staining is evident in the uninfected cells while there is bright RSV-G staining covering the surface of multiple individual and combined cell nuclei in the chamber exposed to RSV.

4.3 ALI Cultured BCi show inconsistent evidence of productive RSV-A2 infection

BCi cells cultured using the ALI method formed a polarised layer of pseudostratified epithelial cells with beating cilia present on the apical surface. H&E staining shows clearly defined cilia at 200x magnification as well as a layer of denser staining cells along the transwell membrane suggestive of basal cells (**Figure 4-4 A**). TEER readings taken every 2 days following removal of the apical media show increasing membrane resistance as the cells differentiated. This is reported as Ohms*CM2 in **Figure 4-4 B**, calculated by multiplying the sample resistance of replicate wells by the surface area of the transwell membrane (0.33cm2).

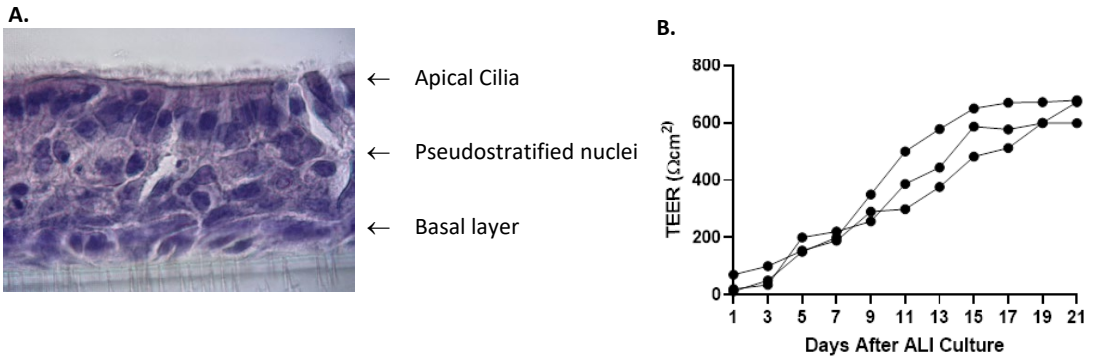


Figure 4-4 BCi ALI culture produced a polarised cell layer of differentiated airway epithelial cells confirmed by H&E staining and TEER

*BCi-NS1.1 cells cultured at the ALI differentiate from basal cells into a polarised layer of pseudostratified airway epithelial cells with apical cilia and mucus production. A. H&E staining of mature BCi ALI. B. TEER readings reported as Ohms*CM2 for average readings taken from n=3 plates of BCi ALI every 2 days. Increasing membrane resistance suggests tight junction formation occurring during ALI culture.*

Initial RSV infection in BCi ALI utilised a 72h infection period with a 2.5ul dose of RSV per insert, equivalent to $1.75x10^5$ PFU. These parameters were based on previous work from the group as well as the relevant literature. Results from PCR analysis of these experiments are shown in **Figure 4-5**. The primary endpoint was *RSV-N* gene amplification over the course of infection, with fold changes calculated for 72hpi relative to 2hpi samples. The four replicate experiments available for analysis of 72h infection resulted in *RSV-N* fold changes of 0.28, 77.37, 18.07 and 10.57 (Replicates 1,2,3 and 5). A fourth replicate experiment was excluded due to loss of the infected 2hpi baseline sample. In the third and fifth replicate experiments, additional samples collected at 48hpi show 3.37- and 1.31-fold increases in *RSV-N*. This variability continued into the *CXCL10* analysis with 25.84, 153.14, 97.92, 5.58-fold increase in expression at 72hpi relative to uninfected 2h samples. The fourth experiment resulted in only 2.13-fold increase in *CXCL10*. Upregulation of *CXCL10* also occurred in infected samples at 2hpi as well as in uninfected samples at 72hpi. There was no significant difference between infected samples and uninfected samples at either time point.

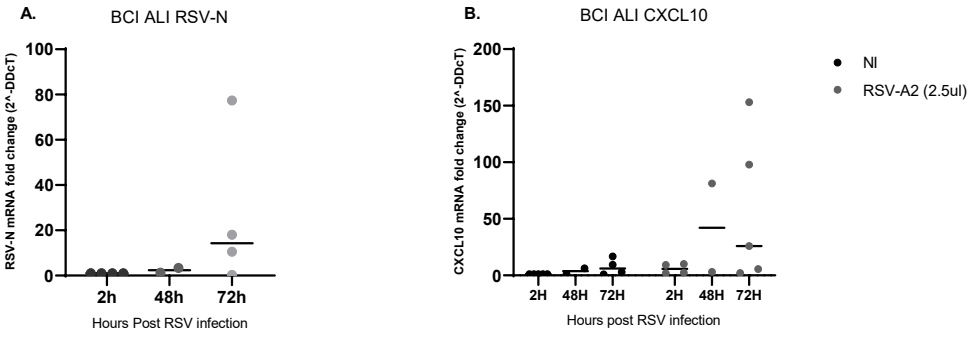


Figure 4-5 BCi-ALI RSV-A2 infection resulted in inconsistent RSV-N amplification and CXCL10 expression.

BCi-ALI were infected apically with 2.5ul (1.75x10^5 PFU) of RSV-A2 in 50ul HBSS for 2h followed by incubation for a further 46 and 70h. RNA samples collected at the 2h, 48h and 72h time points were analysed by RT-qPCR to assess expression of RSV-N and CXCL10 at 72h relative to 2hpi samples (RSV-N) or 2h uninfected controls (CXCL10). There was considerable variability between replicate infections resulting in no statistical significance between conditions. 2h and 72h samples n=5, 48h samples n=2.

The experiment was repeated using BCi-SA ALI with an additional 5ul (3.5x10^5 PFU) RSV-A2 infection, the result of qPCR relative expression of *RSV-N* and *CXCL10* for these experiments are shown in **Figure 4-6**. These cells were less convincingly infected than the larger airway BCi ALI; *RSV-N* was mildly increased (2.28-fold) at 72hpi in one replicate 5ul infection relative to 2hpi, however the remaining infections resulted in an overall slight reduction in RSV-N expression. *CXCL10* was also not upregulated in these cells at 72hpi relative to 2h controls. There was a slight dose dependent increase in the 2hpi samples, however CXCL10 was not upregulated by approximately 2-fold in both infected and uninfected samples at 72h.

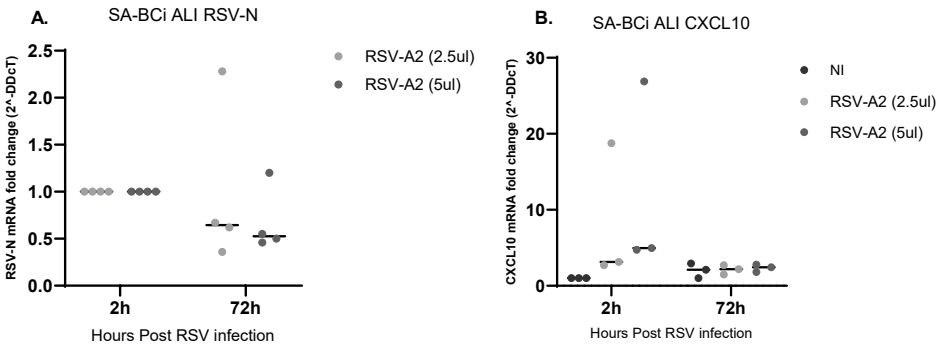
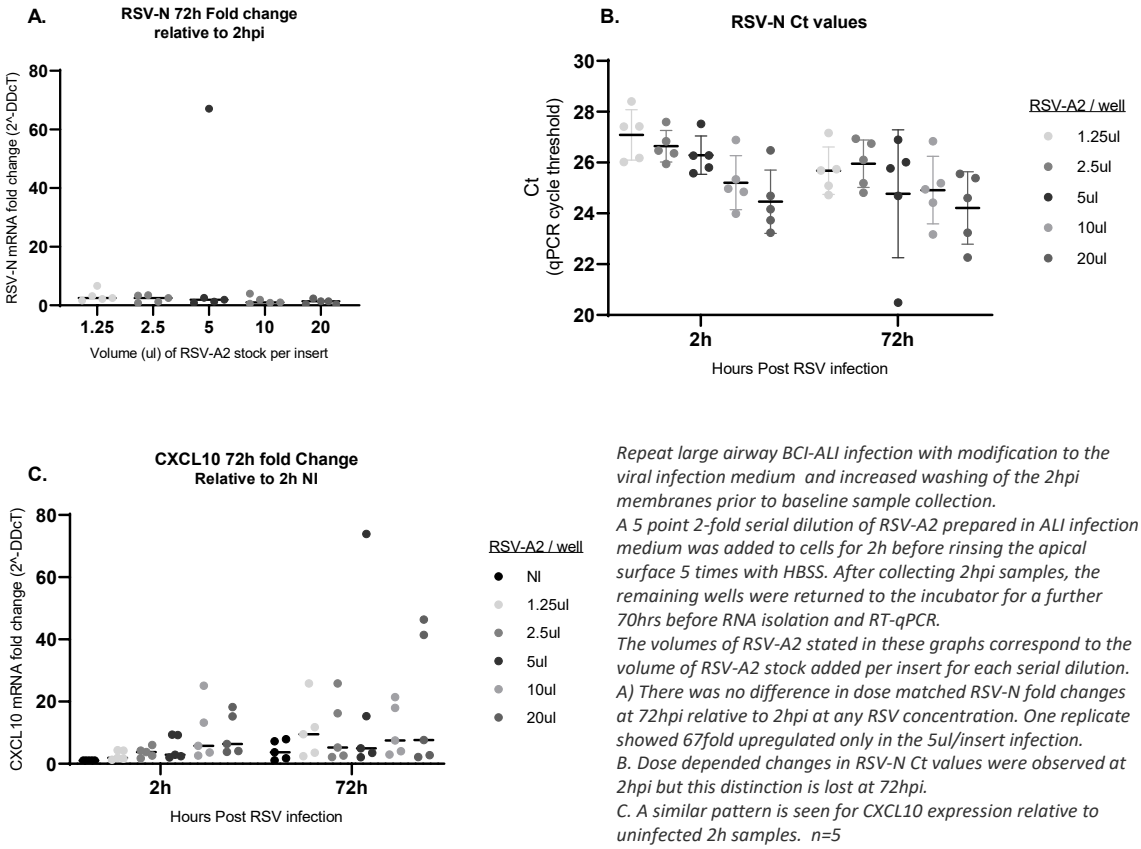


Figure 4-6 SA BCi ALI Do not show changes in RSV-N or CXCL10 expression after 72h RSV-A2 infection.

Small airway BCi ALI were infected with either 2.5ul or 5ul of RSV-A2 (in 50ul HBSS) per insert for 2hours followed by 72h incubation. RNA samples collected 2 and 72 hours post infection (hpi) were reverse transcribed to cDNA and processed for qPCR analysis of RSV-N and CXCL10. A) RSV-N mRNA decreased overall relative to the 2hpi baseline in all but 1 of 4 replicate experiments and only modestly in this sample. B) CXCL10 mRNA increased in an RSV dose dependent manner only at 2hpi. Subsequent 72h samples show less than 2-fold upregulation of CXCL10, equivalent to 72h uninfected samples (N=3)

Given the more productive, yet inconsistent, infection observed with the large airway BCi-ALI, a repeat experiment was devised with modifications made to the protocol to account for potential sources of error. The HBSS inoculum used previously was replaced with serum free ALI infection media, described in methods, and the inserts were rinsed 5 times in HBSS after the 2h infection period before harvesting the 2h samples. Both modifications were based on a similar experiment from the literature. These experiments also included a 5 point 2-fold serial dilution of RSV-A2 including concentrations higher and lower than the 2.5 and 5ul inoculum used previously. *RSV-N* and *CXCL10* mRNA was then measured as before, with each 72h infected sample compared to dose matched 2hpi baseline samples (*RSV-N*) or 2h uninfected controls (*CXCL10*). Results from this analysis are shown in **Figure 4-7.**

There was no difference in RSV-N expression between 72hpi and 2hpi dose matched samples using relative expression calculations. Amplification of *RSV-N* over the course of infection only occurred in one replicate 5ul RSV infection which produced a 67-fold upregulation of *RSV-N* over 72h, all other replicates of this RSV-A2 concentration resulted in fold changes between 0.9 and 1.5. The 1.25ul ($8.75x10^{04}$ PFU) experiments were slightly more consistent yet weakly productive resulting RSV-N upregulation between 1.6 and 6.7-fold. 2.5ul ($1.75x10^{05}$ PFU) infection resulted in 0.8 and 2.5-fold changes, 10ul ($7x10^{05}$ PFU) infection between 0.5 and 3.9-fold and 20ul ($1.4x10^{06}$ PFU) infections between 0.3 and 2.4-fold. There was a dose dependent decrease in *RSV-N* Ct value in the 2hpi samples, however by 72hpi this distinction is lost. A similar pattern was seen for *CXCL10* expression (**Figure 4-7C**), however there was stronger induction of *CXCL10* in some samples, despite low RSV-N expression. The highest induction was seen in one replicate of a 5ul RSV-A2 infection.



Repeat large airway BCI-ALI infection with modification to the viral infection medium and increased washing of the 2hpi membranes prior to baseline sample collection.
A 5 point 2-fold serial dilution of RSV-A2 prepared in ALI infection medium was added to cells for 2h before rinsing the apical surface 5 times with HBSS. After collecting 2hpi samples, the remaining wells were returned to the incubator for a further 70hrs before RNA isolation and RT-qPCR.
The volumes of RSV-A2 stated in these graphs correspond to the volume of RSV-A2 stock added per insert for each serial dilution.
A) There was no difference in dose matched RSV-N fold changes at 72hpi relative to 2hpi at any RSV concentration. One replicate showed 67fold upregulated only in the 5ul/insert infection.
B. Dose depended changes in RSV-N Ct values were observed at 2hpi but this distinction is lost at 72hpi.
C. A similar pattern is seen for CXCL10 expression relative to uninfected 2h samples. n=5

Figure 4-7 Repeat BCI-ALI RSV-A2 infection using a range of concentrations and modified protocol resulted in less evidence of productive infection

4.4 Primary ALI cultures are readily infected with RSV

4.4.1 RSV-G protein visualisation in PBEC cultures

PBEC have been shown to be permissive to RSV infection in the literature so were used as a final experiment to determine whether the poor infection seen in the BCi ALI was cell type dependent. PBEC obtained from resected lung tissue (TL1831) and enriched with bronchial epithelial cells were initially infected as monolayer cultures and prepared for immuno-fluorescent imaging of RSV-G protein. The same low concentrations of RSV-A2 and infection parameters utilised with BCi monolayers were used for this PBEC infection. Red AF-568 staining was evident over the entire

surface of infected PBEC with some nuclei appearing to be covered by the same RSV-G positive membrane. These cells had a less organised appearance and were much larger than the undifferentiated BCi cells. Doubling the concentration of RSV had no noticeable impact on the number of cells with positive staining. **Figure 4-8** displays representative 200x magnification images of duplicate chamber slide compartments exposed RSV-A2 for 72h as well as an uninfected sample.

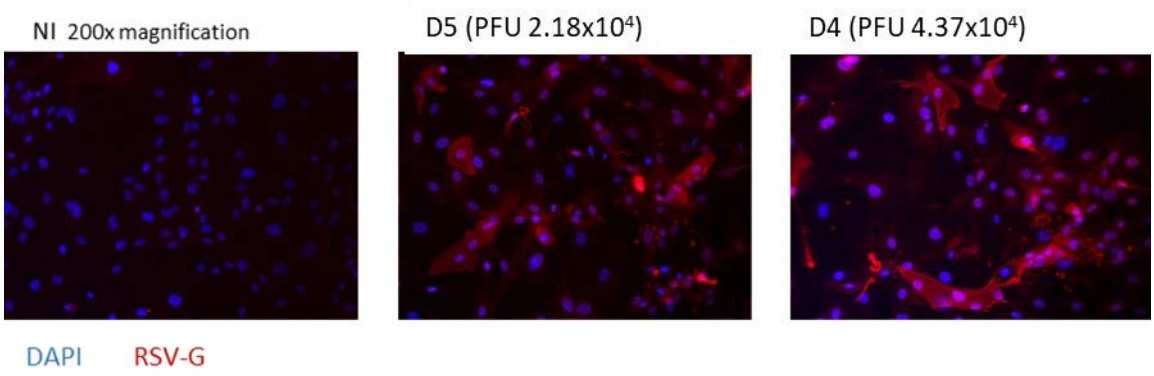


Figure 4-8 PBEC monolayer culture RSV-G immuno-fluorescent imaging

Primary bronchial epithelial cells were cultured in 4 well chamber slides and infected with serial dilutions of RSV-A2. Representative images of D4 and D5 infections show RSV-G staining (red) using AF-568 conjugated secondary antibodies DAPI stained nuclei (blue) at 200x magnification

After ALI differentiation, primary cells from the TL1831 sample produced mucus and had beating cilia on the apical surface. TEER reached an average of 450 Ohms.CM2 21 days after airlift at which point the cells were prepared for infection. H&E staining of the TL1831 ALI following infection with RSV-A2 (2.5ul/ insert, ($1.75x10^{05}$ PFU) shows cilia expression and possible pseudo stratification, however the cell layers are not as defined as the BCi-ALI **Figure 4-9.**

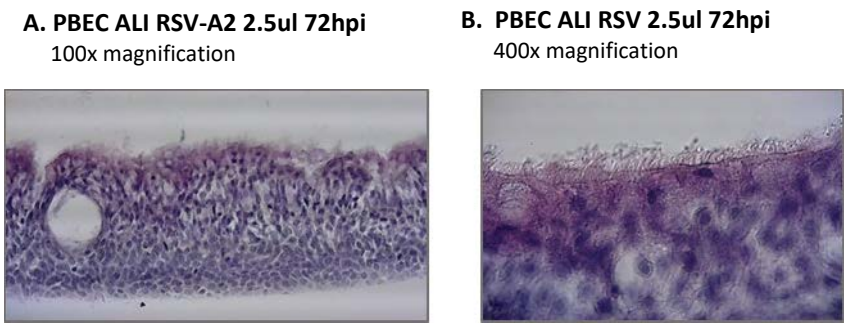


Figure 4-9 Haematoxylin and Eosin staining of TL1831 PBEC ALI

ALI differentiated Primary bronchial epithelial cells were infected with 2.5ul of RSV-A2 per insert for 2 hours followed by a further 70hrs incubation. ALI were fixed in 4% PFA followed by brief storage in 70% ETOH prior to paraffin embedding. The same 5µm section is shown here stained with H&E and imaged at A. 100x and B. 400x magnification. Nuclei are stained blue while cytoplasm and cellular structures including cilia are pink

Immuno-fluorescent staining of the same sample identified RSV-G in the cytoplasm of cells focused on the apical surface of the ALI section (**Figure 4-10 B**). There is also a cluster of nuclei surrounded by RSV-G staining (Arrow). No red staining was present in another section of the same embedded sample used as a control for non-specific binding of the conjugated secondary antibody **Figure 4-10 A**. This sample was exposed to all conditions of the staining protocol with the exception of the RSV-G primary antibody.

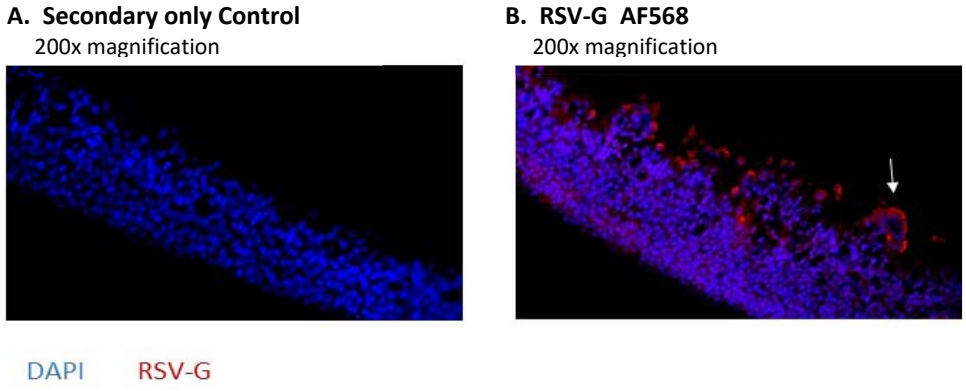


Figure 4-10 PBEC ALI RSV-G Immuno-Fluorescent imaging

Primary bronchial epithelial cells cultured at ALI were infected with RSV-A2(MOI-1) for 2h followed by 72h incubation. Paraffin embedded 5µM sections were stained for RSV-G protein using AF-568 conjugated secondary antibodies (red) and DAPI (nuclei) and TRITC (RSV-G) channels merged to produce image at 200x magnification. A. Primary RSV-G antibody omitted from staining protocol to identify any non-specific secondary antibody binding. B. Fully stained cells.

4.4.2 PBEC ALI antiviral gene expression qPCR

Gene expression was explored in two PBEC samples at the time of writing, the TL1831 sample and an additional TL2022 sample which was not enriched with PBEC before culture (described in methods). TL1831 ALI samples exposed to 2.5, 5 and 10ul of RSV for a total incubation period of 72h were prepared for qRT-PCR. An additional sample, TL2022, was exposed to 2.5ul and 5ul of RSV-A2 for the same duration and using the same experimental conditions. These samples differ only in their donor, the gross architecture of the tissue source and steps taken prior to initial cell culture. *RSV-N* mRNA was increased with all three concentrations of RSV in the TL1831 sample, most prominently in response to 5ul infections **(Figure 4-11. A)**.

The highest concentration of RSV-A2, 10ul, produced a relative decrease in *RSV-N* expression (fold change 17.5) compared to the 2.5ul (fold change 30.84) and 5ul (fold change 60.42) infections. Transcriptional upregulation of *CXCL10* was several hundred times higher in infected samples at 72hpi relative to the 2h uninfected control **(Figure 4-11. B & C)**. RSV-A2 used at 5ul per well produced 895-fold induction of CXCL10 at 72hpi, 2.5ul RSV-A2 resulted in 415-fold upregulation and the highest 10ul RSV-A2 concentration resulted in 666-fold upregulation. *CXCL10* expression also shows a dose dependent increase in the 2hpi samples ranging from 3.6-12.9-fold upregulation in the 2.5, 5 and 10ul RSV-A2 infections respectively. In uninfected samples, *CXCL10* was only 4.3-fold higher at 72h relative to 2h.

TL2022 ALI samples produced more modest 18.36-fold upregulation of *RSV-N* at 72hpi with 2.5ul infections. This decreases to 11.09-fold upregulation with the more concentrated 5ul infection, however, Ct values for *RSV-N* were lower for the higher concentration infections at each time point. *CXCL10* expression does show a dose dependent increase at 72hpi, with 169.48 and 210.64-fold upregulation for 2.5 and 5ul RSV infection respectively.

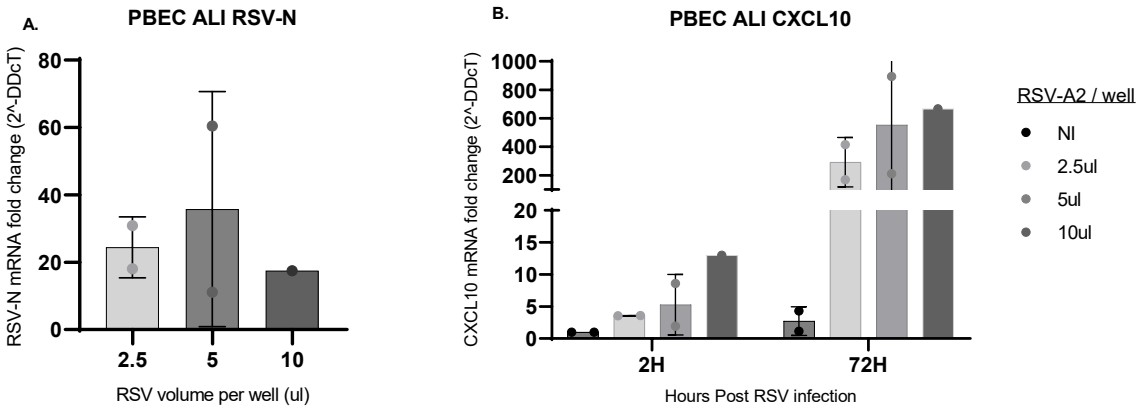


Figure 4-11 Primary ALI RSV-N and Antiviral gene expression by qPCR

Primary bronchial epithelial cells cultured at ALI were infected with RSV for 2h followed by a further 70h incubation period. A. Relative expression of RSV-N at 72hpi calculated in reference to matched 2hpi samples. B. CXCL10 expression was calculated relative to 2h uninfected samples. The CXCL10 gene plot used a split y axis to show all data points more clearly. RSV-A2 2.5ul and 5ul infection N=2, RSV-A2 10ul infection n=1.

4.5 Discussion

This project aimed to explore the use of the BCiNS1.1 basal cell line as an ALI model of RSV infection. These cells have shown good differentiation capacity both in the introductory paper and in this project. Defined and beating cilia were consistently observed on the apical surface within a week of airlift and can be seen in paraffin embedded sections stained with H&E. Despite this, the RSV-A2 infections were not repeatable either across replicate experiments or within an individual plate. The new stocks of RSV-A2 purchased for use in this work appear to be viable as they produced evidence of cytopathic effect and possible syncytia formation in monolayer cultures and *RSV-N* was amplified over time in some ALI infections. Immuno-fluorescent staining of RSV-A2 infected BCi monolayers showed RSV-G staining coating the entire surface of discreet patches of cells. There was also evidence of multinucleated cells and possible cell lysis. Although additional staining of cellular structures would be needed to confirm, these multinucleated cells could be evidence of syncytia

formation which is known to occur during RSV infection *in vitro* (54). Undifferentiated BCi are basal cells isolated from the large airways of a healthy non-smoker. Basal cells have been shown to be permissive to RSV infection *in vitro* (134), and this may occur in the infected lung following damage to overlaying epithelial cell layer. Other immortalised airway epithelial cell monolayer cultures have been infected with RSV and produced similar staining patterns to those shown in this project. Low MOI infection of 2D cultures of the BEAS-2B simian virus 40 (SV40)-transformed bronchial epithelial cell line produced small foci of RSV infection at 48hpi with no visible difference in the number of cells infected with RSV at MOI 0.1 or MOI 0.3 at 24 or 48hpi suggesting that only certain cells are initially infected (135).

While no studies have yet shown RSV infection in BCi ALI cultures, the work conducted in this project suggests that RSV-A2 infection can occur, at least in the large airway BCi cell line. Differences in susceptibility to RSV have been observed in ALI cultures of PBEC from different airway compartments. In one study (54), well differentiated cultures of primary paediatric bronchial epithelial cells were slightly more susceptible to infection than well differentiated nasal cells from the same person. Another study (136), reported reduced RSV-A2 susceptibility in small airway cells (SmallAir: primary small-airway human airway epithelial cells) than in upper airway cells (MucilAir: primary nasal or bronchial human airway epithelial cells). Both cell types were infected however replication was found to be slower in the small airway cells although cytokine production was comparable. In this project, both *CXCL10* and *RSV-N* were consistently not upregulated in SA-BCi ALI despite the cells initially showing a response to infection with CXCL10 upregulation at 2hpi.

RSV is thought to preferentially infect ciliated epithelial cells from the apical surface and to enter the cells either by direct fusion with the target cell plasma membrane or via receptor mediated endocytosis. The rate of infection is therefore heavily dependent on expression of viral attachment proteins in the culture model. Unlike IAV which can utilise NA to spread between cells, a study using well differentiated primary airway epithelial cells from the upper airways suggests that RSV spread

between cells occurs at the apical surface and depends on cilia beating to move viral particles through the airway mucus/fluid layer (137). The cells became more permissive to infection as they differentiated into a ciliated phenotype and infection was limited to ciliated cells even when the cell layer was compromised. The differentiation state of the ALI used in this project was confirmed using TEER and visual confirmation of cilia and mucus production. H&E staining was only conducted on the samples shown however and additional staining of markers of differentiation, such goblet cells or tight junctions, was not carried out. In the absence of these experiments, it is possible that the variation observed in this work could be due to inadequate differentiation of samples.

Increasing the number of apical washes from 2 to 5 could have improved the protocol by ensuring that all unbound RSV was removed prior to 2hpi sample collection and continued incubation of the 72 or 48h samples. Some replicates of each experiment, especially BCi-1 which resulted in a 0.83-fold decrease in RSV-N, did have lower than average RSV-N Ct values in the 2hpi samples. If RSV-A2 was not fully removed at this point it would both artificially reduce the calculated relative expression and could increase the antiviral response of the cells, as measured here by *CXCL10*, due to prolonged exposure to RSV. In this case, the dose curve experiment resulted in even poorer evidence of RSV infection; however, it is unclear if this is due to the increased washes or the use of media instead of HBSS for the viral inoculum. In retrospect, the proprietary PneumaCult ALI 10x supplement used in the ALI infection media could contain substances which would inhibit infection when applied with the viral inoculum.

The PBEC ALI infections utilised the original protocol and resulted in stronger evidence of RSV infection suggesting that differences in RSV infection are likely cell based. There was considerably stronger *CXCL10* induction in PBEC ALI than in the BCi-ALI in both donor samples as well as more convincing RSV-N amplification, especially in the TL1831 cells. RSV-G protein was also shown coating the surface of apical cells in RSV infected TL1831 ALI. The differences in *RSV-N* and *CXCL10* fold change data between cells from each donor could represent donor variation in RSV susceptibility.

The 2hpi samples from both donors produced a linear decrease of approximately 1 Ct per doubled RSV concentration as would be expected. TL1831 ALI infection resulted doubling of *RSV-N* and *CXCL10* fold changes between 2ul and 5ul infection at 72hpi but a relative reduction in fold change in response to the highest 10ul infection. This could indicate human error during sample or experiment preparation or could indicate decreased cell viability or reduced viral replication with this concentration. Excessive concentrations of RSV have been shown to impair infection due to an elevated number of short interfering particles, defective fragments of RSV which prevent productive replication (138). The TL2022 infections resulted in lower overall fold induction of both *RSV-N* and *CXCL10* than was seen for the TL1831 samples. There was also a decline in *RSV-N* fold change in response to 5ul RSV infection compared to 2.5ul infection. In this case, there was still a dose dependent increase in the antiviral response of the cells as measured by *CXCL10* expression. If these are true findings, the TL2022 samples may be less susceptible to RSV-A2 infection than the TL1381 samples, possibly due to cellular characteristics or differentiation state. TL1831 ALI were generated from a sample containing visible airways which were scraped to enhance the PBEC concentration of the initial cultures. This was not possible for the TL2022 tissue sample which had no visible airway structures. Greater information about the phenotype of these cells would need to be obtained by continuing the imaging experiments to show adequate differentiation as well as entry receptor expression and RSV localisation in the cells.

The larger airway BCI ALI show evidence of productive infection in some replicates but poor consistency meant that there were no significant differences in *RSV-N* or *CXCL10* expression between 72hpi and 2h samples overall. If the ALI model can be optimised to improve consistency, it holds promise as a reliable model of RSV infection. Confirmation of ALI differentiation as well as the health of the cells prior to infection would be a necessary next step to achieve this goal. Whilst TEER is a useful proxy for differentiation, it only measures changes in electrical resistance over the membrane and cannot give an indication of full differentiation or cell viability. In this work, epithelial barrier integrity stabilises at around day 15 post airlift, and the cells were used for

infection at day 21 to day 28. No confirmation of differentiation state beyond that time frame, or evaluation of changes which might occur during the window of infection have been carried out. In addition to viability, general cell stress could impact the infection results. *CXCL10* was used as a measure of antiviral response in these experiments as it is known to be produced early into infection with RSV. However, *CXCL10* is also mildly upregulated in uninfected cells over the same time frame. To confirm whether changes in *CXCL10* are the result of productive infection, inactive RSV-A2 should be included as an additional control.

Chapter 5 Discussion and Future Work

This project was successfully able to identify differences in T cell related gene expression in response to the viruses discussed, supporting the main hypothesis of this work. However, there were limitations to the conclusions which can be drawn from this work alone. Most importantly, transcriptomics analysis from these culture models cannot provide insight on the wider inflammatory response in the airways. Inflammatory cells will introduce additional signalling molecules in response to those released by the infected epithelium providing feedback which could change the overall presentation in natural infection. Additionally, the PBEC- T cell gene signature identified in this work contains genes with varied functionality. Determining whether their expression in response to each virus will impact T cell signalling from PBEC would require more direct measurement of this interaction. Transcription alone does not always correlate with translation, meaning that the differences identified here could also be lost due to variations in protein expression. With more time, the differences between H1N1 and H3N2 responses would receive more attention. Specifically, the GSE47961 dataset is associated with two other subseries containing additional repeats of the same experiments as well as H1N1 infection at MOI-1. Including these arrays could improve the significance of the findings and add additional information about the impact of MOI on the H1N1 data. Similarly, super series GSE32140, which includes the GSE32138 subseries, contains datasets exploring PBMC responses to RSV and IAV which would also be explored.

While this was not shown to be reliable by the time of submission, with additional work, the BCi-NS1.1 cell line could be a suitable candidate for RSV infection modelling. ALI differentiated models of infection are increasingly used in the literature however no one has shown BCI-ALI infection with RSV. Future work on this model will need to improve consistency in ALI generation, both in terms of the passage number used for seeding, the length of time between airlift and infection. The former was not possible to control during this project and the latter was maintained within a reasonable

window however, as the DE analysis has shown, days can make a big difference in the antiviral and developmental state of ALI differentiated cells. Additional steps to confirm validity of the model would be to monitor differentiation state more closely, for example through measurement of cilia gene expression or imaging of differentiation markers (cilia, goblet cells, tight junctions). With a stable baseline model, the number, differentiation state and viability of cells per well could be more confidently assumed, allowing for more accurate and repeatable infection using MOI.

The design would be further improved by including controls for non-productive infection. *CXCL10* qPCR demonstrated that low levels of upregulation occur in uninfected cells as well as infected. While the difference should be considerable with productive infection as seen with the preliminary PBEC work, it would be beneficial to control for non-specific activation of the cells. Non- productive infection may also result in elevated CXCL10. To control for this, UV-inactivated RSV should be included with future efforts to develop the BCi cell line as an RSV infection model.

Despite its limitations, the comparative transcriptomics analysis did produce a large data set with numerous avenues for further investigation. While this project focused on T cell related functions, with more time, contrasting the differences in metabolic pathway modulation and ECM degradation for example could identify key drivers of lung damage and differences in the dynamics of disease resolution during infection. The differences between H1N1 and H3N2 would be further explored as a priority, Finally, comparing the wider gene expression profiles seen here with data sets generated from human challenge studies could add additional validity to the findings.

Appendix

A1. R CODE: Agilent Microarray Data Import and Targets File creation

```
#Load required packages
library(GEOquery)
library(limma)
library (dplyr)


## Get metadata from GEO and assign group variables
# load series and platform data from GEO

gset <- getGEO("GSE32138", GSEMatrix =TRUE, AnnotGPL=TRUE)
if (length(gset) > 1) idx <- grep("GPL6480", attr(gset, "names")) else idx <- 1
gset <- gset[[idx]]

# make proper column names on feature variables to match toptable
fvarLabels(gset) <- make.names(fvarLabels(gset))


#Extract fData and pData
FeatureAnn<- fData(gset)
colnames(FeatureAnn)
FeatureAnn <- FeatureAnn[, c(1:4, 10,16:22)] ## subset useful columns

Phenodata <- pData(gset)
colnames(Phenodata)
Phenodata2 <- Phenodata[, c(1,2,8)]

# Create new columns containing grouping information. Matching to terms in existing columns
Phenodata2$Group <- ifelse(grepl("for Influenza",Phenodata2$title ),"Mock.IAV",
                           ifelse(grepl("for RSV", Phenodata2$title),"Mock.RSV",
                           ifelse (grepl("RSV", Phenodata2$title), "RSV",
                                   "IAV")))

### Make targets file
##Format = SamlpleNo, FileName, Condition

SampleNo<- 1:16

FileName<- list.files("GSE32138_RAW")


Targets<- data.frame(SampleNo, FileName)

# Create new columns with useful metadata
Targets$Samples<- Phenodata2$geo_accession
Targets$Condition <- Phenodata2$Group


##Use LIMMA's read.maimages function to load data into an EListRaw class object
x <- read.maimages(Targets, path = "GSE32138_RAW", source="agilent",green.only=TRUE)

dim(x)  ## 45015 probes and 16 arrays
```

```
library(GEOquery)
library(limma)
library(dplyr)

# load series and platform data from GEO

gset <- getGEO("GSE47961", GSEMatrix =TRUE, AnnotGPL=TRUE)
length(gset)
gset <- gset[[1]]


#Extract phenotype data as a dataframe and view what sample information is available
pd <- pData(gset)
names(pd)

# Filter for chosen columns using their indecies and rename columns
pd2 <- pd[, c(1,2,38,40,42)]

colnames(pd2)<- c("Title", "Accession", "BioRep", "Infection", "TimePoint")

# Create a new column with grouping info "Infection_TimePoint"
pd2$Group <- paste(pd2$Infection,pd2$TimePoint, sep= "_")

# Extract feature data and reduce to chosen columns
Feat<- fData(gset)
colnames(Feat)
Feat <- Feat[, c(1:4, 10,16:22)]


### Make a Targets File for data upload
##Format = SamlpleNo, FileName, Group, SampleID
### Raw data files were selected and downloaded manually from NCBI GEO supplementary files.
### Current WD contains raw file folder "GSE47961_RAW"

SampleNo<- 1:88

FileName<- list.files("GSE47961_RAW")

Targets<- data.frame(SampleNo, FileName)

Targets$Group <- pd2$Group
Targets$SampleID <- pd2$Accession
Targets$Virus <- pd2$Infection
Targets$Time <- pd2$TimePoint

colnames(Targets) <- c("SampleNo", "FileName", "Group", "SampleID", "Virus")

## Change group labels from eg "icSARS_0h" to "SARS_0h"

Targets$Group <- gsub("ic", "", as.character(Targets$Group))
```

A2. R CODE: LIMMA Normalisation and QC Plots GSE47961

```
### Use LIMMA's read.maimages function to load data into an EListRAW object.
x <- read.maimages(Targets, path = "GSE47961_RAW", source="agilent",green.only=TRUE)

dim(x)  ## 45015 probes and 88 arrays

EXPR <- x$E ## Have a look

## Rename samples/arrays with unique titles eg SARS_24hr_3, extracting info from pd2

Targets$Title <- pd2$Title
Targets$Title <- gsub("SHAE003_", "", as.character(Targets$Title))
Targets$Title <- gsub("ic", "", as.character(Targets$Title))

colnames(x$E) <- Targets$Title
colnames(x$Eb) <- Targets$Title ## dont forget background data

x$targets <- Targets ## updating targets file


#### Normalise and background correct in one step, using the backround and control data provided

x_NEQC <- neqc(x, status= x$genes$ControlType, negctrl=-1,
               regular=0)

### neqc produces an EList-class or matrix object containing normalized log2 intensities,
# with control probes removed

dim(x_NEQC) # 43376 obs, control probes removed


### Average replicate spots
x_NEQC_aveRep <- avereps(x_NEQC, ID=x_NEQC$genes$ProbeName)
```

```
### Data QC ###

## Create a factor of categorical grouping variable for use in colouring plots
GrF <- factor(Targets$Group, levels = unique(Targets$Group))
GrF2 <- factor(unique(GrF))

ViF <- factor(Targets$Virus,levels = unique(Targets$Virus))
relevel (ViF, ref = "Mock")
ViF2 <- factor(unique(ViF))
relevel (ViF2, ref = "Mock")


# Specify colour palette
palette(c("#1B9E77", "#7570B3", "#E7298A", "#E6AB02", "#D95F02",
          "#66A61E", "#A6761D", "#666666"))


## boxplots of Raw and norm data
boxplot(log2(x$E), boxwex=0.6, notch=T, outline=FALSE, las=2, col=ViF ,
        main = "GSE47961 RAW", ylab= "Log2 Intensity", names = Targets$Title,
        cex.axis =0.75)
legend("topright", levels(ViF2), fill=palette(), bty="0", bg= "white")


boxplot((x_NEQC_aveRep$E), boxwex=0.6, notch=T, outline=FALSE, las=2, col=ViF ,
        main = "GSE47961 NEQC", ylab= "Log2 Intensity", names = Targets$Title,
        cex.axis =0.75)
legend("topright", levels(ViF2), fill=palette(), bty="0", bg= "white")


## Hierarchical clustering analysis
dst <- dist(t(x_NEQC_aveRep$E))
hc <- hclust(dst, method="complete")

plot(hc, label=GrF, main="GSE47961 NEQC", cex= 0.6, cex.main= 0.8, hang= -1)


## Multidimensional scaling
plotMDS(x_NEQC_aveRep, main= "MDS GSE47961 NEQC" , labels = Targets$Time, xlab = "Dim1", ylab = "Dim2",
        col = as.integer(ViF))
legend("topright",col= as.integer(ViF2), legend =levels(ViF2) ,pch = 16, cex=0.8, inset=c(0.02))
```

```
## Many unlabeled genes in provided feature data, improve using clusterprofiler
#gene annotation
library(clusterProfiler)
library(hgug4112a.db)

my_probes <-x_NEQC_aveRep$genes$ProbeName
Symbol <-bitr(my_probes,fromType = "PROBEID",toType = "SYMBOL",OrgDb = hgug4112a.db,drop = FALSE)
EntrezID <-bitr(my_probes,fromType = "PROBEID",toType = "ENTREZID",OrgDb = hgug4112a.db,drop = FALSE)
## 27.62% of input gene IDs are fail to map... Expected.

x_NEQC_Final <- x_NEQC_aveRep
x_NEQC_Final$genes<-cbind(x_NEQC_aveRep$genes,Symbol,EntrezID)

## have a look:
x_NEQC_Final_genes <- x_NEQC_Final$genes ## dim 41000 14
colnames(x_NEQC_Final_genes)

# Remove excess columns (control and row cols not needed anymore, compare orgininal and new gene symbols)
x_NEQC_Final_genes <- x_NEQC_Final_genes[,c("ProbeName", "GeneName","PROBEID","SYMBOL","ENTREZID")]

## New annotation has more mapped genes, NAs introduced for unmapped instead of repeating probe ID

x_NEQC_Final_genes<- x_NEQC_Final_genes[!is.na(x_NEQC_Final_genes$SYMBOL), ]

dim(x_NEQC_Final_genes) ## 29677     5

x_NEQC_Final$genes<- x_NEQC_Final_genes

## make sure expr data matches
probes <- x_NEQC_Final$genes %>%
  pull (ProbeName)

x_NEQC_Final$E <- x_NEQC_Final$E [probes, ]

dim(x_NEQC_Final$E) # 29677     88

## No substantial change to MDS
plotMDS(x_NEQC_Final,main= "MDS GSE47961 NEQC" , labels = Targets$Time, xlab = "Dim1", ylab = "Dim2",
        col = as.integer(ViF))
legend("topright",col= as.integer(ViF2), legend =levels(ViF2) ,pch = 16, cex=0.8, inset=c(0.02))
```

A3. R CODE: Agilent Microarray Differential Expression Analysis

```
###### Differential Expression Analysis #######
### Build the design matrix for the linear modelling function ###

design <- model.matrix(~0 + f)

colnames(design) <- levels(f)
design

## Fit the linear model to the design matrix
fit <- lmFit(y.ave_Filt, design)

### Define contrasts
contrast.matrix<-makeContrasts (RSV = RSV - Mock.RSV,
                                IAV = IAV - Mock.IAV,
                                levels = design)
contrast.matrix



### Apply contrast matrix to the modeled data and compute statistics for the data.
fit2 <- contrasts.fit(fit, contrast.matrix)
fit2 <- eBayes(fit2)
```

```
### Subset data and remove unlabeled probes and duplicated genes
OutputIAV <- topTable(fit2,coef = 2, adjust.method ="BH", sort.by="B", number = Inf)
dim(OutputIAV) ### 41000 22

OutputIAV <- subset(OutputIAV, select= c ("ProbeName", "Gene.symbol","Gene.title",

                                          "logFC", "P.Value","adj.P.Val", "B" ))

colnames(OutputIAV)<- c("ID", "Gene.symbol","Gene.title",

                        "IAV.logFC", "IAV.P.Val","IAV.adj.P.Val", "IAV.B")


OutputIAV_Symbol <- OutputIAV$Gene.symbol != ""

OutputIAV <- OutputIAV[OutputIAV_Symbol, ]  ## subset rows using logical vector
dim(OutputIAV)  ## 30723   7

OutputIAV <- OutputIAV %>%
  group_by(Gene.symbol) %>%
  arrange(desc(IAV.logFC)) %>%
  dplyr::slice(1)
dim(OutputIAV)   #19565   7

# Remove grouping factor
OutputIAV <- data.frame(OutputIAV)
```

```
##### time Matched DE analysis ########
## Define experiment using Group variable, SARS_24h, H1N1_6h etc

design <- model.matrix(~0+ GrF)
colnames(design) <- levels(GrF)


## Fit the linear model to the design matrix, define contrsts, apply contrasts to model
fit <- lmFit(x_NEQC_Final, design)


contrastMat <- makeContrasts (H1N1.24h = H1N1_6hr - Mock_6hr,
                              H1N1.12h = H1N1_12hr - Mock_12hr,
                              H1N1.18h = H1N1_18hr - Mock_18hr,
                              H1N1.24h = H1N1_24hr - Mock_24hr,
                              H1N1.324h = H1N1_36hr - Mock_36hr,
                              H1N1.48h = H1N1_48hr - Mock_48hr,
                              SARS.24h = SARS_24hr - Mock_24hr,
                              SARS.48h = SARS_48hr - Mock_48hr,
                              SARS.240h = SARS_60hr - Mock_60hr,
                              SARS.72h = SARS_72hr - Mock_72hr,
                              SARS.84h = SARS_84hr - Mock_84hr,
                              SARS.824h = SARS_924hr - Mock_924hr,
                              levels = design)


fit2 <- contrasts.fit(fit, contrastMat)

# compute statistics and table of top significant genes
fit2 <- eBayes(fit2)

## Extract results for each contrast using fit2 coefficients
## Repeated for each coef
colnames(fit2)

H1N1.24h_TT <- topTable(fit2, coef= 4, adjust="fdr", sort.by="B", number= Inf)
colnames(H1N1.6h_TT)

H1N1.24h_TT <- subset(H1N1.24h_TT, select=c("PROBEID","SYMBOL","logFC","AveExpr",
                                            "adj.P.Val","P.Value","B"))


## Collapse to gene level using aggregate, taking mean average of all test stats

H1N1.24h_Agg <- aggregate(cbind(H1N1.24h_TT$logFC, H1N1.24h_TT$P.Value, H1N1.24h_TT$adj.P.Val),
                          list(H1N1.24h_TT$SYMBOL), mean)

colnames(H1N1.24h_Agg) <- c("SYMBOL", "logFC","P.Value", "adj.P.Val")

dim(H1N1.24h_Agg) ### 18812     4
```

A4. R CODE: GSE47961 maSigPro analysis

```
########## GSE47961 MaSigPro ######

library(maSigPro)

## Using Targets file and NEQC normalised expression data fro Time matched analysis
## Create separate expression objects for each virus

idx <- which(Targets$Virus == "H1N1" | Targets$Virus == "Mock")
H1N1_neqc <- x_NEQC_Final[,idx]

dim(H1N1_neqc) ## 41000    60

## Need only time matched mock samples
# Get H1N1 time points
H1N1_Time <- Targets %>%
  filter (Virus == "H1N1") %>%
  pull (Time)

H1N1_Time <- unique (H1N1_Time)


H1N1_neqcT <- H1N1_neqc$targets

idx <- which(H1N1_neqcT$Time %in% H1N1_Time, )

## filter new data set
H1N1_keep <- H1N1_neqcT[idx , ]
H1N1_keep2 <- H1N1_neqc[, idx ]


## Create matrix of norm expression values

H1N1_E <- H1N1_keep2$E
dim(H1N1_E)  ## 41000    48

library(WGCNA)
options(stringsAsFactors = FALSE)
## collapseRows needs data to be numeric, extract rowIDS (probe names) and groupIDs (gene symbols) as vectors then create a nea matrix
## with only the numeric data. collapseRows asks for the group and row information,

rowGroup <- H1N1_neqc$genes$SYMBOL
rowID <- H1N1_neqc$genes$PROBEID

H1N1_coll <- collapseRows(datET=H1N1_E,
                          rowGroup= rowGroup,
                          rowID= rowID,
                          method="MaxMean")   ## warning, should remove genes without symbols/nas before running. removed by function anyway

H1N1_Collapse <- H1N1_coll$datETcollapsed  ## 18,812 rows, labeled with gene symbols, looks good, highest expression value retained

str(H1N1_Collapse)
```

```
## Create experiment design matrix defining each infection

eDesignH <- data.frame (H1N1_keep2$targets$Time, H1N1_keep2$targets$Title)
colnames(eDesignH) <- c("Time", "Title")


eDesignH$Time <- gsub("hr", "", as.character(eDesignH$Time))
eDesignH$Time <- as.numeric (eDesignH$Time)  ## eDesign needs to be numeric!, Time col was character

Replicates <- c(rep(1:7, each = 4), rep(8:10, each = 3), rep(11:11, each = 2), rep(12:14, each = 3))

eDesignH$Replicate <- Replicates

eDesignH$Control <- ifelse(grepl("Mock",eDesignH$Title),1,0)

eDesignH$H1N1 <- ifelse(grepl("H1N1",eDesignH$Title),1,0)

rownames(eDesignH) <- eDesignH$Title

# Remove Title col
eDesignH <- eDesignH[ , -2]
str(eDesignH)

eDesignH <- as.matrix(eDesignH)
str(eDesignH)

identical(colnames(H1N1_Collapse), rownames(eDesignH))

#Create a regression matrix for the full regression model. 7 time points per infection group so 6 DF

designH <- make.design.matrix(eDesignH, degree = 6)

## p.vector computes a regression fit for each gene
## min.obs = genes with less than this number of true numerical values will be excluded.
##  Minimum value to estimate the model is (degree+1)xGroups+1. Default is 6.
## My design specifies 6 degrees (No. of time variables 7 - 1) and 2 groups? (Control, SARS/H1N1) (6+1)x (2+1) = 21


fitH <- p.vector(H1N1_Collapse, designH, Q = 0.05, MT.adjust = "BH", min.obs = 21)

## p.vector() returns a list of values:
## > fit$i # returns the number of significant genes

fitH$i  ## 13346


######## T.fit selects the best regression model for each gene using stepwise regression ######

tstepH <- T.fit(fitH, step.method = "backward", min.obs = 21, alfa = 0.05)
```

```
### Obtain lists of SIGNIFICANT GENES FOR EACH EXPERIMENTAL GROUP
##  R-squared value restricts the model selection to genes with good fit: Here the models must be greater than 0.7.

sigsH <- get.siggenes(tstepH, rsq = 0.7, vars = "groups") ## creates list of significant genes

### VISUALISE SIGNIFICANT GENES AS CLUSTERS

H1N1_k9 <- see.genes(sigsH$sig.genes$H1N1vsControl, show.fit = T, dis =designH$dis,
                     cluster.method="hclust" ,cluster.data = 1, k = 9)

## Pull expression and significance data for each data set. The replicates of each time point are not averaged
H1N1_Sig_Profiles <- sigsH$sig.genes$H1N1vsControl$sig.profiles
H1N1_P <- sigsH$sig.genes$H1N1vsControl$sig.pvalues

## Extract expression values and P values for genes in each cluster
H1N1_k9_C1 <- names(which(H1N1_k9$cut==1))
H1N1_k9_C2 <- names(which(H1N1_k9$cut==2))
```

A5. R CODE: ClusterProfiler ORA

```
############### ClusterProfiler ORA of GSE47961 H1N1 up regulated genes ############

library(clusterProfiler)
library(enrichplot)
# we use ggplot2 to add x axis labels (ex: ridgeplot)
library(ggplot2)
library(hgug4112a.db)
library(DOSE)

# reading in log2FC data from limma
H1N1 = read.csv("H1N1.24h_Sig.log2FC.csv", header=TRUE)
H.clusters <- read.csv("H1N1.Clusters.csv", header = TRUE)

## MappSYMBOLS to ENTREZID, needed for goENRICH
H1.Symbols <-H1N1$SYMBOL
H1.Entrez <-bitr(H1.Symbols ,fromType = "SYMBOL",toType = "ENTREZID",OrgDb = hgug4112a.db,drop = FALSE)

H1N1.2 <- merge(H1N1, H1.Entrez, by = "SYMBOL")

## create vectors of ordered LFC, named with Entres ID
H1.UP.EZlist <- H1N1.2$logFC
names(H1.UP.EZlist)<- H1N1.2$ENTREZID
H1.UP.EZlist <- sort(H1.UP.EZlist, decreasing = TRUE) ##1538

H1.UP.EZlist<- na.omit(H1.UP.EZlist) # none
head(H1.UP.EZlist)

### Repeat for Cluster gene lists
H1.C5 <- H.clusters[, "C.5"]
H1.C5 <- na.omit(H1.C5)

H1.C5_Sub <- H1N1.2[H1N1.2$SYMBOL %in%  H1.C5, ]

H1.C5_LFC <- H1.C5_Sub$logFC
names(H1.C5_LFC) <- H1.C5_Sub$ENTREZID
H1.C5_LFC <- sort (H1.C5_LFC, decreasing = TRUE)

### Run ORA analysis using enrichGO ######

go_C.5_2 <- enrichGO(gene = names(H1.C5_LFC), ## named LFC vector for cluster 5 genes
                     universe = names(H1.UP.EZlist), ## named lfc vector for all upregulated genes
                     OrgDb = hgug4112a.db,
                     keyType = 'ENTREZID',
                     readable = T,
                     ont = "BP",
                     pvalueCutoff = 0.05,
                     qvalueCutoff = 0.1)

barplot(go_C.5_2,
        drop = TRUE,
        showCategory = 10,
        title = "H1N1 Cluster 5 GO Biological Pathways",
        font.size = 12)
```

```
## T cell pathway imaging ##
## Create dataframes with all genes with adjP <0.05 in H1N1 24h data
## Subset for up and down reg genes

H1N1.24hALL = read.csv("H1N1.24h_All_Sig_LFC.csv", header=TRUE)
colnames(H1N1.24hALL) <- c("SYMBOL", "24h.LFC", "24h_AdjP")

H1.24hA.Sym <- H1N1.24hALL$SYMBOL
H1.24A.Entrez <-bitr(H1.24hA.Sym ,fromType = "SYMBOL",toType = "ENTREZID",OrgDb = hgug4112a.db,drop = FALSE)
H1N1.24hALL.2 <- merge(H1N1.24hALL, H1.24A.Entrez, by = "SYMBOL")

H1N1.24_SigUP <- subset(H1N1.24hALL.2, H1N1.24hALL.2$`24h.LFC` >= 1 )
H1N1.24_sigDown <- subset(H1N1.24hALL.2, H1N1.24hALL.2$`24h.LFC` <= -1)

## Create named LFC vector of all adjP <0.05 genes in H1N1 24h data
H1.24H_ALL_EZlist <- H1N1.24hALL.2$`24h.LFC`
names(H1.24H_ALL_EZlist) <- H1N1.24hALL.2$ENTREZID
H1.24H_ALL_EZlist <- sort(H1.24H_ALL_EZlist, decreasing = TRUE)
head(H1.24H_ALL_EZlist)

## Create named LFC vector of all adjP <0.05 genes with log2>1 in H1N1 24h data
H1.24h.Up.EZlist <- H1N1.24_SigUP$`24h.LFC`
names(H1.24h.Up.EZlist)<- H1N1.24_SigUP$ENTREZID
H1.24h.Up.EZlist <- sort(H1.24h.Up.EZlist, decreasing = TRUE)

H1.24h.Up.EZlist<- na.omit(H1.24h.Up.EZlist) # none
head(H1.24h.Up.EZlist)


go_H1.24h.UP <- enrichGO(gene = names(H1.24h.Up.EZlist),    ### 24h up reg (lfc >= 1), from below original list
                         universe = names(H1.24H_ALL_EZlist), ## original 24h gene list sig (adjP < 0.05) filter only
                         OrgDb = hgug4112a.db,
                         keyType = 'ENTREZID',
                         readable = T,
                         ont = "BP",
                         pvalueCutoff = 0.05,
                         qvalueCutoff = 0.1)


barplot(go_H1.24h.UP,
        drop = TRUE,
        showCategory = 10,
        title = "H1N1 24h GO Biological Pathways",
        font.size = 12)
```

```
## to access genes in the 4th enriched term Defence response to Virus
defRespVirus <- go_H1.24h.UP@result$geneID[4]  # creates character vector, symbols split with /

defRespVirus2 <- unlist(strsplit(defRespVirus, split = "/"))


## T cell terms defined using a list of pre research GO:BP IDs
TcellIDs <- c("GO:0010818","GO:0072678", "GO:0042110", "GO:0050868", "GO:0050870", "GO:0019882")

## extract all IDs from 24h up enrichGO result
H1.24UP_IDs <- go_H1.24h.UP@result$ID[]

# Extract indexes for the location of these IDs in the result
TcellSubIND <- which(H1.24UP_IDs %in% TcellIDs)

## Subset result for using this vector. Generates a DF with resukts for chosen terms
TcellSub <- go_H1.24h.UP[TcellSubIND]

## ggplot attempt at barplot visualisation
ggplot(TcellSub2,aes(x=Description,y=Count,fill=p.adjust)) +
  geom_col(position="dodge",width=0.8,size=0.8) +
  coord_flip()+
  scale_fill_gradient(low = "red", high = "blue") +
  ggtitle("H1N1 24h upregulated T cell pathways") +
  theme(axis.text.y = element_text(size = 12))
```

References

1. Chatzis O, Darbre S, Pasquier J, Meylan P, Manuel O, Aubert JD, et al. Burden of severe RSV disease among immunocompromised children and adults: a 10 year retrospective study. BMC Infect Dis [Internet]. BioMed Central; 2018 [cited 2018 Aug 17];18(1):111. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29510663

2. Tan WC, Xiang X, Qiu D, Ng TP, Lam SF, Hegele RG. Epidemiology of respiratory viruses in patients hospitalized with near-fatal asthma, acute exacerbations of asthma, or chronic obstructive pulmonary disease. Am J Med [Internet]. 2003 Sep [cited 2018 Aug 17];115(4):272–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12967691

3. Carvajal LA, Pérez CP. Epidemiology of Respiratory Infections. Pediatr Respir Dis [Internet]. Nature Publishing Group; 2020 [cited 2021 Oct 11];263. Available from: /pmc/articles/PMC7120591/

4. WHO. Influenza (Seasonal) [Internet]. 2018 [cited 2021 Aug 4]. Available from: https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal)

5. Falsey AR, McElhaney JE, Beran J, van Essen GA, Duval X, Esen M, et al. Respiratory Syncytial Virus and Other Respiratory Viral Infections in Older Adults With Moderate to Severe Influenza-like Illness. J Infect Dis [Internet]. Narnia; 2014 Jun 15 [cited 2019 Mar 23];209(12):1873–81. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jit839

6. Walsh EE, Peterson DR, Kalkanoglu AE, Lee FE-H, Falsey AR. Viral Shedding and Immune Responses to Respiratory Syncytial Virus Infection in Older Adults. J Infect Dis [Internet]. Oxford University Press; 2013 May 1 [cited 2018 Aug 17];207(9):1424–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23382572

7. Ackerson B, Tseng HF, Sy LS, Solano Z, Slezak J, Luo Y, et al. Severe Morbidity and Mortality Associated With Respiratory Syncytial Virus Versus Influenza Infection in Hospitalized Older Adults. Clin Infect Dis [Internet]. Oxford Academic; 2019 Jul 2 [cited 2021 Nov 20];69(2):197–203. Available from: https://academic.oup.com/cid/article/69/2/197/5193205

8. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet [Internet]. Lancet Publishing Group; 2017 Sep 2 [cited 2021 Jan 11];390(10098):946–58. Available from: https://pubmed.ncbi.nlm.nih.gov/28689664/

9. Price RHM, Graham C, Ramalingam S. Association between viral seasonality and meteorological factors. Sci Rep [Internet]. Nature Publishing Group; 2019 Dec 1 [cited 2021 Dec 5];9(1). Available from: /pmc/articles/PMC6353886/

10. World Health Organisation. Weekly epidemiological update on Coronavirus Disease (COVID-19) [Internet]. 2023 [cited 2023 May 24]. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---11-may-2023

11. World Health Organisation Eastern Mediterranean Regional Office. MERS situation update, December 2019 [Internet]. 2019 [cited 2021 Nov 2]. Available from: http://www.emro.who.int/pandemic-epidemic-diseases/mers-cov/mers-situation-update-december-2019.html

12. Chan-Yeung M, Xu R. SARS: epidemiology. Respirology [Internet]. Respirology; 2003 Nov [cited 2021 Nov 2];8 Suppl(Suppl 1). Available from: https://pubmed.ncbi.nlm.nih.gov/15018127/

13. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung S, Hayashi K, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). Int J Infect Dis [Internet]. 2020 [cited 2020 Jul 4];94:154–5. Available from: https://doi.org/10.1016/j.ijid.2020.03.020

14. Byambasuren O, Cardona M, Bell K, Clark J, McLaws ML, Glasziou P. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: Systematic review and meta-analysis. J Assoc Med Microbiol Infect Dis Canada [Internet]. University of Toronto Press; 2020 Dec 1 [cited 2021 Nov 7];5(4):223–34. Available from: https://jammi.utpjournals.press/doi/abs/10.3138/jammi-2020-0030

15. Alvarez-Munoz S, Upegui-Porras N, Gomez AP, Ramirez-Nieto G. Key Factors That Enable the Pandemic Potential of RNA Viruses and Inter-Species Transmission: A Systematic Review. Viruses [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2021 Apr 1 [cited 2021 Nov 15];13(4). Available from: /pmc/articles/PMC8063802/

16. Centers for Disease Control and Prevention National Center for Immunization and Respiratory Diseases (NCIRD). 2009 H1N1 Pandemic (H1N1pdm09 virus) | Pandemic Influenza (Flu) [Internet]. cdc.gov website. 2019 [cited 2021 Dec 3]. Available from: https://www.cdc.gov/flu/pandemic-resources/2009-h1n1-pandemic.html

17. Taubenberger JK, Morens DM. Influenza: The Once and Future Pandemic. Public Health Rep [Internet]. SAGE Publications; 2010 [cited 2021 Dec 3];125(Suppl 3):16. Available from: /pmc/articles/PMC2862331/

18. WHO. Past pandemics [Internet]. World Health Organisation Europe. Denmark: World Health Organization; 2021 [cited 2021 Dec 3]. Available from: https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/pandemic-influenza/past-pandemics

19. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nat 2015 5237559 [Internet]. Nature Publishing Group; 2015 Jun 8 [cited 2023 May 17];523(7559):217–20. Available from: https://www.nature.com/articles/nature14460

20. Cohen LE, Spiro DJ, Viboud C. Projecting the SARS-CoV-2 transition from pandemicity to endemicity: Epidemiological and immunological considerations. PLoS Pathog [Internet]. PLOS; 2022 Jun 1 [cited 2023 May 22];18(6). Available from: /pmc/articles/PMC9246171/

21. Public Health England. Guidance on use of antiviral agents for the treatment and prophylaxis of seasonal influenza. 2021.

22. Malik S, Ahmad T, Muhammad K, Waheed Y. Respiratory Syncytial Virus Infection: Treatments and Clinical Management. Vaccines [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2023 Feb 1 [cited 2023 May 17];11(2). Available from: /pmc/articles/PMC9962240/

23. Forbester JL, Humphreys IR. Genetic influences on viral-induced cytokine responses in the lung [Internet]. Vol. 14, Mucosal Immunology. Springer Nature; 2020 [cited 2021 Jan 13]. p. 14–25. Available from: https://doi.org/10.1038/s41385-020-00355-6

24. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Møller R, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. Cell [Internet]. Cell Press; 2020 May 28 [cited 2020 Jul 1];181(5):1036–1045.e9. Available from: https://doi.org/10.1016/j.cell.2020.04.026

25. Acharya D, Liu G, Gack MU. Dysregulation of type I interferon responses in COVID-19. Nat Rev Immunol.

26. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, et al. Dysregulation of immune response in patients with COVID-19 in Wuhan, China. Clin Infect Dis An Off Publ Infect Dis Soc Am [Internet]. Oxford University Press; 2020 Aug 1 [cited 2021 Aug 29];71(15):762–8. Available from: /pmc/articles/PMC7108125/?report=abstract

27. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol 2021 197 [Internet]. Nature Publishing Group; 2021 Jun 1 [cited 2021 Nov 21];19(7):409–24. Available from: https://www.nature.com/articles/s41579-021-00573-0

28. Jafri HS, Wu X, Makari D, Henrickson KJ. Distribution of Respiratory Syncytial Virus Subtypes A and B Among Infants Presenting to the Emergency Department With Lower Respiratory Tract Infection or Apnea. Pediatr Infect Dis J [Internet]. 2013 Apr [cited 2019 Mar 24];32(4):335–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23337904

29. Ruzin A, Pastula ST, Levin-Sparenberg E, Jiang X, Fryzek J, Tovchigrechko A, et al. Characterization of circulating RSV strains among subjects in the OUTSMART-RSV surveillance program during the 2016-17 winter viral season in the United States. PLoS One [Internet]. Public Library of Science; 2018 [cited 2019 Mar 24];13(7):e0200319. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30040837

30. Laham FR, Mansbach JM, Piedra PA, Hasegawa K, Sullivan AF, Espinola JA, et al. Clinical Profiles of Respiratory Syncytial Virus Subtypes A and B among Children Hospitalized with Bronchiolitis. Pediatr Infect Dis J [Internet]. NIH Public Access; 2017 Aug 1 [cited 2021 Aug 30];36(8):808. Available from: /pmc/articles/PMC5556381/

31. Griffiths C, Drews SJ, Marchant DJ. Respiratory Syncytial Virus: Infection, Detection, and New Options for Prevention and Treatment. Clin Microbiol Rev [Internet]. American Society for Microbiology Journals; 2017 Jan 1 [cited 2018 Dec 11];30(1):277–319. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27903593

32. Afonso CL, Amarasinghe GK, Bányai K, Bào Y, Basler CF, Bavari S, et al. Taxonomy of the order Mononegavirales: update 2016. Arch Virol [Internet]. Springer Vienna; 2016 Aug 23 [cited 2019 Mar 23];161(8):2351–60. Available from: http://link.springer.com/10.1007/s00705-016-2880-1

33. Battles MB, McLellan JS. Respiratory syncytial virus entry and how to block it. Nat Rev Microbiol. Nature Publishing Group; 2019 Apr 1;17(4):233–45.

34. Leung NHL. Transmissibility and transmission of respiratory viruses. Nat Rev Microbiol 2021 198 [Internet]. Nature Publishing Group; 2021 Mar 22 [cited 2021 Nov 22];19(8):528–45. Available from: https://www.nature.com/articles/s41579-021-00535-6

35. Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, Hegele RG. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. Nat Med 2011 179 [Internet]. Nature Publishing Group; 2011 Aug 14 [cited 2021 Nov 2];17(9):1132–5. Available from: https://www.nature.com/articles/nm.2444

36. Krzyzaniak MA, Zumstein MT, Gerez JA, Picotti P, Helenius A. Host Cell Entry of Respiratory Syncytial Virus Involves Macropinocytosis Followed by Proteolytic Activation of the F Protein. PLoS Pathog [Internet]. Public Library of Science; 2013 Apr [cited 2021 Nov 15];9(4):1003309. Available from: /pmc/articles/PMC3623752/

37. Jang J, Bae SE. Comparative Co-Evolution Analysis Between the HA and NA Genes of Influenza A Virus. Virology (Auckl) [Internet]. SAGE Publications; 2018 Jul 19 [cited 2021 Nov 18];9. Available from: /pmc/articles/PMC6053862/

38. Zhou X, McElhaney JE. Age-related changes in memory and effector T cells responding to influenza A/H3N2 and pandemic A/H1N1 strains in humans. Vaccine [Internet]. NIH Public Access; 2011 Mar 3 [cited 2018 Jan 15];29(11):2169–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21353149

39. Wan H, Perez DR. Quail carry sialic acid receptors compatible with binding of avian and human influenza viruses. Virology [Internet]. Virology; 2006 Mar 15 [cited 2021 Nov 18];346(2):278–86. Available from: https://pubmed.ncbi.nlm.nih.gov/16325879/

40. Dou D, Revol R, Östbye H, Wang H, Daniels R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. Front Immunol [Internet]. Frontiers Media SA; 2018 Jul 20 [cited 2021 Oct 24];9(JUL):1. Available from: /pmc/articles/PMC6062596/

41. Sempere Borau M, Stertz S. Entry of influenza A virus into host cells — recent progress and remaining challenges. Curr Opin Virol. Elsevier; 2021 Jun 1;48:23–9.

42. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet [Internet]. 2020 [cited 2020 May 12];395:565. Available from: https://www.ncbi.

43. Wang M-Y, Zhao R, Gao L-J, Gao X-F, Wang D-P, Cao J-M. SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. Front Cell Infect Microbiol. Frontiers; 2020 Nov 25;0:724.

44. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. Virol J [Internet]. BioMed Central Ltd.; 2019 May 27 [cited 2021 Nov 21];16(1):69–69. Available from: https://europepmc.org/articles/PMC6537279

45. Murgolo N, Therien AG, Howell B, Klein D, Koeplinger K, Lieberman LA, et al. SARS-CoV-2 tropism, entry, replication, and propagation: Considerations for drug discovery and development. PLOS Pathog [Internet]. 2021;17(2):e1009225. Available from: https://dx.plos.org/10.1371/journal.ppat.1009225

46. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science [Internet]. Science; 2020 Mar 13 [cited 2021 Nov 21];367(6483):1260–3. Available from: https://pubmed.ncbi.nlm.nih.gov/32075877/

47. Taubenberger JK, Morens DM. The Pathology of Influenza Virus Infections. Annu Rev Pathol [Internet]. NIH Public Access; 2008 [cited 2023 May 16];3:499. Available from: /pmc/articles/PMC2504709/

48. Kuiken T, Taubenberger JK. Pathology of human influenza revisited. Vaccine [Internet]. NIH Public Access; 2008 Sep 12 [cited 2021 Nov 12];26(Suppl 4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2605683/

49. Bagga B, Woods CW, Veldman TH, Gilbert A, Mann A, Balaratnam G, et al. Comparing inluenza and RSV viral and disease dynamics in experimentally infected adults predicts clinical effectiveness of RSV antivirals. Antivir Ther. 2013;18(6):785–91.

50. Johnson JE, Gonzales RA, Olson SJ, Wright PF, Graham BS. The histopathology of fatal untreated human respiratory syncytial virus infection. Mod Pathol 2007 201 [Internet]. Nature Publishing Group; 2006 Nov 24 [cited 2021 Nov 15];20(1):108–19. Available from: https://www.nature.com/articles/3800725

51. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a

descriptive study. Lancet [Internet]. 2020 [cited 2020 Jun 22];395:507–13. Available from: https://doi.org/10.1016/

52. Z W, JM M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA [Internet]. JAMA; 2020 Apr 7 [cited 2021 Aug 30];323(13):1239–42. Available from: https://pubmed.ncbi.nlm.nih.gov/32091533/

53. Pickles RJ, DeVincenzo J. RSV and its propensity for causing bronchiolitis. J Pathol [Internet]. 2015;235(2):266. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25302625

54. Guo-Parke H, Canning P, Douglas I, Villenave R, Heaney LG, Coyle P V., et al. Relative Respiratory Syncytial Virus Cytopathogenesis in Upper and Lower Respiratory Tract Epithelium. Am J Respir Crit Care Med Vol [Internet]. American Thoracic Society; 2013 Oct 1 [cited 2021 Oct 18];188(7):842–51. Available from: www.atsjournals.org

55. Miura TA. Respiratory Epithelial Cells as Master Communicators during Viral Infections [Internet]. Vol. 6, Current Clinical Microbiology Reports. Springer; 2019 [cited 2021 Jul 2]. p. 10–7. Available from: https://doi.org/10.1007/s40588-019-0111-8

56. Tam A, Wadsworth S, Dorscheid D, Paul Man SF, Sin DD. The airway epithelium: more than just a structural barrier. Ther Adv Respir Dis [Internet]. 2011 [cited 2021 Oct 12];5(4):255–73. Available from: http://www.sagepub.co.uk/

57. Tilley AE, Walters MS, Shaykhiev R, Crystal RG. Cilia Dysfunction in Lung Disease. Annu Rev Physiol [Internet]. NIH Public Access; 2015 Feb 1 [cited 2021 Aug 9];77:379. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4465242/

58. Hsieh MH, Beirag N, Murugaiah V, Chou YC, Kuo WS, Kao HF, et al. Human Surfactant Protein D Binds Spike Protein and Acts as an Entry Inhibitor of SARS-CoV-2 Pseudotyped Viral Particles. Front Immunol. Frontiers Media S.A.; 2021 May 14;12:1613.

59. Chuquimia OD, Petursdottir DH, Periolo N, Ferńndez C. Alveolar Epithelial Cells Are Critical in Protection of the Respiratory Tract by Secretion of Factors Able To Modulate the Activity of Pulmonary Macrophages and Directly Control Bacterial Growth. Infect Immun [Internet]. American Society for Microbiology (ASM); 2013 Jan [cited 2021 Dec 6];81(1):381. Available from: /pmc/articles/PMC3536158/

60. Fehrenbach H. Alveolar epithelial type II cell: defender of the alveolus revisited. Respir Res [Internet]. BioMed Central; 2001 [cited 2021 Nov 28];2(1):33. Available from: /pmc/articles/PMC59567/

61. Ye J, Zhang B, Xu J, Chang Q, McNutt MA, Korteweg C, et al. Molecular Pathology in the Lungs of Severe Acute Respiratory Syndrome Patients. Am J Pathol [Internet]. American Society for Investigative Pathology; 2007 [cited 2021 Dec 7];170(2):538. Available from: /pmc/articles/PMC1851867/

62. I I, F Y, B M, M W, E F. Toll-like receptor expression and induction of type I and type III interferons in primary airway epithelial cells. J Virol [Internet]. J Virol; 2013 Mar 15 [cited 2021 Oct 21];87(6):3261–70. Available from: https://pubmed.ncbi.nlm.nih.gov/23302870/

63. Jensen S, Thomsen AR. Sensing of RNA Viruses: a Review of Innate Immune Receptors Involved in Recognizing RNA Virus Invasion. J Virol [Internet]. American Society for Microbiology (ASM); 2012 Mar 15 [cited 2021 Dec 7];86(6):2900. Available from: /pmc/articles/PMC3302314/

64. Melkamu T, Squillace D, Kita H, O'Grady SM. Regulation of TLR2 expression and function in

human airway epithelial cells. J Membr Biol [Internet]. Springer; 2009 May 10 [cited 2021 Dec 7];229(2):101–13. Available from: https://link.springer.com/article/10.1007/s00232-009-9175-3

65. Guillot L, Le Goffic R, Bloch S, Escriou N, Akira S, Chignard M, et al. Involvement of Toll-like Receptor 3 in the Immune Response of Lung Epithelial Cells to Double-stranded RNA and Influenza A Virus *. J Biol Chem [Internet]. Elsevier; 2005 Feb 18 [cited 2021 Dec 7];280(7):5571–80. Available from: http://www.jbc.org/article/S0021925819630590/fulltext

66. Mielcarska MB, Bossowska-Nowicka M, Toka FN. Cell Surface Expression of Endosomal Toll-Like Receptors—A Necessity or a Superfluous Duplication? Front Immunol. Frontiers Media S.A.; 2021 Feb 1;11:3652.

67. Levy DE, Marié IJ, Durbin JE. Induction and Function of Type I and III Interferon in Response to Viral Infection. Curr Opin Virol [Internet]. NIH Public Access; 2011 [cited 2021 Dec 7];1(6):476. Available from: /pmc/articles/PMC3272644/

68. Sommereyns C, Paul S, Staeheli P, Michiels T. IFN-Lambda (IFN-λ) Is Expressed in a Tissue-Dependent Fashion and Primarily Acts on Epithelial Cells In Vivo. PLoS Pathog [Internet]. Public Library of Science; 2008 Mar [cited 2021 Dec 7];4(3):1000017. Available from: /pmc/articles/PMC2265414/

69. See H, Wark P. Innate immune response to viral infection of the lungs. Paediatr Respir Rev. 2008;9.

70. Zhang W, Zhang L, Zan Y, Du N, Yang Y, Tien P. Human respiratory syncytial virus infection is inhibited by IFN-induced transmembrane proteins. J Gen Virol [Internet]. J Gen Virol; 2015 Jan 1 [cited 2021 Nov 14];96(Pt 1):170–82. Available from: https://pubmed.ncbi.nlm.nih.gov/25228491/

71. Glaser L, Coulter PJ, Shields M, Touzelet O, Power UF, Broadbent L. Airway epithelial derived cytokines and chemokines and their role in the immune response to respiratory syncytial virus infection. Pathogens [Internet]. MDPI AG; 2019 Sep 1 [cited 2021 May 23];8(3). Available from: /pmc/articles/PMC6789711/

72. Chan MCW, Cheung CY, Chui WH, Tsao GSW, Nicholls JM, Chan YO, et al. Proinflammatory cytokine responses induced by influenza A (H5N1) viruses in primary human alveolar and bronchial epithelial cells. Respir Res [Internet]. BioMed Central; 2005 Nov 11 [cited 2021 Nov 28];6(1):135. Available from: /pmc/articles/PMC1318487/

73. Thwaites RS, Coates M, Ito K, Ghazaly M, Feather C, Abdulla F, et al. Reduced nasal viral load and IFN responses in infants with respiratory syncytial virus bronchiolitis and respiratory failure. Am J Respir Crit Care Med [Internet]. American Thoracic Society; 2018 Oct 15 [cited 2023 May 6];198(8):1074–84. Available from: www.atsjournals.org.

74. Mueller SN, Rouse BT. Immune responses to viruses. Clin Immunol [Internet]. Elsevier; 2008 [cited 2021 Dec 3];421. Available from: /pmc/articles/PMC7151814/

75. Roumanes D, Falsey AR, Quataert S, Secor-Socha S, Lee -H, Yang H, et al. T-Cell Responses in Adults During Natural Respiratory Syncytial Virus Infection. J Infect Dis [Internet]. Oxford University Press; 2018 Jul 2 [cited 2018 Aug 17];418(3):218. Available from: https://academic.oup.com/jid/article/218/3/418/5039135

76. Pishesha N, Harmand TJ, Ploegh HL. A guide to antigen processing and presentation. Nat Rev Immunol 2022 2212 [Internet]. Nature Publishing Group; 2022 Apr 13 [cited 2023 May 8];22(12):751–64. Available from: https://www.nature.com/articles/s41577-022-00707-2

77. Yates A, Bergmann C, Leo Van Hemmen J, Stark J, Callard R. Cytokine-modulated Regulation of Helper T Cell Populations. J Theor Biol. Academic Press; 2000 Oct 21;206(4):539–60.

78. Wan Z, Zhou Z, Liu Y, Lai Y, Luo Y, Peng X, et al. Regulatory T cells and T helper 17 cells in viral infection. Scand J Immunol [Internet]. John Wiley & Sons, Ltd; 2020 May 1 [cited 2021 Nov 18];91(5):e12873. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/sji.12873

79. Jozwik A, Habibi MS, Paras A, Zhu J, Guvenel A, Dhariwal J, et al. RSV-specific airway resident memory CD8+ T cells and differential disease severity after experimental human infection. Nat Commun [Internet]. Nature Publishing Group; 2015 Dec 21 [cited 2018 Aug 18];6(1):10224. Available from: http://www.nature.com/articles/ncomms10224

80. Sridhar S, Begom S, Bermingham A, Hoschler K, Adamson W, Carman W, et al. Cellular immune correlates of protection against symptomatic pandemic influenza. Nat Med 2013 1910 [Internet]. Nature Publishing Group; 2013 Sep 22 [cited 2021 Dec 1];19(10):1305–12. Available from: https://www.nature.com/articles/nm.3350

81. Wilkinson TM, Li CKF, Chui CSC, Huang AKY, Perkins M, Liebner JC, et al. Preexisting influenza-specific CD4 + T cells correlate with disease protection against influenza challenge in humans. Nat Med [Internet]. Nature Publishing Group; 2012;18(2):274–80. Available from: http://dx.doi.org/10.1038/nm.2612

82. Sanchez AM, Zhu J, Huang X, Yang Y. The Development and Function of Memory Regulatory T Cells After Acute Viral Infections. J Immunol [Internet]. NIH Public Access; 2012 Sep 15 [cited 2021 Dec 5];189(6):2805. Available from: /pmc/articles/PMC3436958/

83. Bert N Le, Tan AT, Kunasegaran K, Tham CYL, Hafezi M, Chia A, et al. Different pattern of pre-existing SARS-COV-2 specific T cell immunity in SARS-recovered and uninfected individuals. bioRxiv [Internet]. 2020 [cited 2020 Jul 14];2020.05.26.115832. Available from: http://biorxiv.org/content/early/2020/05/27/2020.05.26.115832.abstract

84. Yang L, Peng H, Zhaoling Z, Li G, Huang Z, Zhao Z, et al. Persistent memory CD4+ and CD8+ T-cell responses in recovered severe acute respiratory syndrome (SARS) patients to SARS coronavirus M antigen. J Gen Virol [Internet]. NIH Public Access; 2007 Oct [cited 2020 Jun 23];88(10):2740–8. Available from: /pmc/articles/PMC2362397/?report=abstract

85. Uhl LFK, Gérard A. Modes of Communication between T Cells and Relevance for Immune Responses. Int J Mol Sci [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2020 Apr 2 [cited 2023 May 25];21(8). Available from: /pmc/articles/PMC7215318/

86. Heidema J, Rossen JWA, Lukens M V., Ketel MS, Scheltens E, Kranendonk MEG, et al. Dynamics of Human Respiratory Virus-Specific CD8+ T Cell Responses in Blood and Airways during Episodes of Common Cold. J Immunol [Internet]. American Association of Immunologists; 2008 Oct 15 [cited 2023 May 8];181(8):5551–9. Available from: https://journals.aai.org/jimmunol/article/181/8/5551/78605/Dynamics-of-Human-Respiratory-Virus-Specific-CD8-T

87. Welliver TP, Reed JL, Welliver RC. Respiratory syncytial virus and influenza virus infections: observations from tissues of fatal infant cases. Pediatr Infect Dis J [Internet]. 2008 Oct;27(10 Suppl):S92—6. Available from: https://doi.org/10.1097/INF.0b013e318168b706

88. Siefker DT, Vu L, You D, McBride A, Taylor R, Jones TL, et al. Respiratory syncytial virus disease severity is associated with distinct CD8+ T-cell profiles. Am J Respir Crit Care Med [Internet]. American Thoracic Society; 2020 Feb 1 [cited 2021 Nov 15];201(3):325–34. Available from: www.atsjournals.org.

89. Lee FEH, Walsh EE, Falsey AR, Lumb ME, Okam N V., Liu N, et al. Human Infant Respiratory Syncytial Virus (RSV)—Specific Type 1 and 2 Cytokine Responses Ex Vivo during Primary RSV Infection. J Infect Dis [Internet]. Oxford Academic; 2007 Jun 15 [cited 2021 Dec 6];195(12):1779–88. Available from: https://academic.oup.com/jid/article/195/12/1779/809374

90. Rahimzadeh M, Naderi N. Toward an understanding of regulatory T cells in COVID-19: A systematic review. J Med Virol [Internet]. John Wiley & Sons, Ltd; 2021 Jul 1 [cited 2021 Dec 5];93(7):4167–81. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.26891

91. De Biasi S, Meschiari M, Gibellini L, Bellinazzi C, Borella R, Fidanza L, et al. Marked T cell activation, senescence, exhaustion and skewing towards TH17 in patients with COVID-19 pneumonia. Nat Commun [Internet]. Nature Publishing Group; 2020 Jul 6 [cited 2020 Jul 12];11(1):3434. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32632085

92. Diao B, Wang C, Tan Y, Chen X, Liu Y, Ning L, et al. Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). Front Immunol [Internet]. Frontiers Media S.A.; 2020 May 1 [cited 2020 Jul 1];11:827. Available from: https://www.frontiersin.org/article/10.3389/fimmu.2020.00827/full

93. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat Biotechnol 2020 [Internet]. Nature Publishing Group; 2020 Jun 26 [cited 2021 Aug 7];38(8):970–9. Available from: https://www.nature.com/articles/s41587-020-0602-4

94. De Jong EC, Smits HH, Kapsenberg ML. Dendritic cell-mediated T cell polarization. Springer Semin Immunopathol [Internet]. Springer; 2005 Jan 14 [cited 2021 Dec 7];26(3):289–307. Available from: https://link.springer.com/article/10.1007/s00281-004-0167-1

95. Liu YJ. TSLP in Epithelial Cell and Dendritic Cell Cross Talk. Adv Immunol [Internet]. NIH Public Access; 2009 [cited 2021 Nov 22];101(C):1. Available from: /pmc/articles/PMC3645262/

96. Mennechet FJD, Uzé G. Interferon-lambda-treated dendritic cells specifically induce proliferation of FOXP3-expressing suppressor T cells. Blood [Internet]. Blood; 2006 Jun 1 [cited 2021 Nov 28];107(11):4417–23. Available from: https://pubmed.ncbi.nlm.nih.gov/16478884/

97. Reibman J, Hsu Y, Chen LC, Bleck B, Gordon T. Airway Epithelial Cells Release MIP-3α/CCL20 in Response to Cytokines and Ambient Particulate Matter. https://doi.org/101165/rcmb2002-0095OC [Internet]. American Thoracic Society; 2012 Dec 20 [cited 2021 Oct 16];28(6):648–54. Available from: www.atsjournals.org

98. Rahmatpanah F, Agrawal S, Jaiswal N, Nguyen HM, McClelland M, Agrawal A. Airway epithelial cells prime plasmacytoid dendritic cells to respond to pathogens via secretion of growth factors. Mucosal Immunol [Internet]. Nature Publishing Group; 2019 Jan 1 [cited 2021 Oct 16];12(1):77. Available from: /pmc/articles/PMC6301110/

99. Lund FE, Partida-Sánchez S, Lee BO, Kusser KL, Hartson L, Hogan RJ, et al. Lymphotoxin-α-Deficient Mice Make Delayed, But Effective, T and B Cell Responses to Influenza. J Immunol [Internet]. American Association of Immunologists; 2002 Nov 1 [cited 2022 Sep 19];169(9):5236–43. Available from: https://www.jimmunol.org/content/169/9/5236

100. Olmos S, Stukes S, Ernst JD. Ectopic activation of M. tuberculosis-specific CD4+ T cells in lungs of CCR7−/− mice. J Immunol [Internet]. NIH Public Access; 2010 Jan 1 [cited 2022 Sep 19];184(2):895. Available from: /pmc/articles/PMC2879893/

101. Sallusto F, Lanzavecchia A, MacKay CR. Chemokines and chemokine receptors in T-cell priming and Th1/Th2-mediated responses. Immunol Today. Elsevier Current Trends; 1998 Dec 1;19(12):568–74.

102. Schleimer RP, Kato A, Kern R, Kuperman D, Avila PC. Epithelium: At the interface of innate and adaptive immune responses [Internet]. Vol. 120, Journal of Allergy and Clinical Immunology. NIH Public Access; 2007 [cited 2021 May 3]. p. 1279–84. Available from: /pmc/articles/PMC2810155/

103. Wang Z, Zhu L, Nguyen THO, Wan Y, Sant S, Quiñones-Parra SM, et al. Clonally diverse CD38+HLA-DR+CD8+ T cells persist during fatal H7N9 disease. Nat Commun [Internet]. Nature Publishing Group; 2018 Dec 1 [cited 2020 Jun 21];9(1):1–12. Available from: https://www.nature.com/articles/s41467-018-03243-7

104. Telcian AG, Laza-Stanca V, Edwards MR, Harker JA, Wang H, Bartlett NW, et al. RSV-Induced Bronchial Epithelial Cell PD-L1 Expression Inhibits CD8+ T Cell Nonspecific Antiviral Activity. J Infect Dis [Internet]. Oxford University Press; 2011 Jan 1 [cited 2021 Dec 6];203(1):85. Available from: /pmc/articles/PMC3086441/

105. Ioannidis I, McNally B, Willette M, Peeples ME, Chaussabel D, Durbin JE, et al. Plasticity and Virus Specificity of the Airway Epithelial Cell Immune Response during Respiratory Virus Infection. J Virol [Internet]. American Society for Microbiology; 2012 May 15 [cited 2020 Dec 6];86(10):5422–36. Available from: https://pubmed.ncbi.nlm.nih.gov/22398282/

106. Mitchell HD, Eisfeld AJ, Sims AC, McDermott JE, Matzke MM, Webb-Robertson BJM, et al. A Network Integration Approach to Predict Conserved Regulators Related to Pathogenicity of Influenza and SARS-CoV Respiratory Viruses. PLoS One [Internet]. Public Library of Science; 2013 Jul 25 [cited 2021 Mar 4];8(7). Available from: /pmc/articles/PMC3723910/

107. Bolstad BM. Pre-Processing DNA Microarray Data.

108. S D, PS M. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics [Internet]. Bioinformatics; 2007 Jul 15 [cited 2021 Nov 2];23(14):1846–7. Available from: https://pubmed.ncbi.nlm.nih.gov/17496320/

109. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res [Internet]. 2015 Apr 20 [cited 2018 Apr 30];43(7):e47–e47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25605792

110. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov [Internet]. Cell Press; 2021 Aug 28 [cited 2022 Oct 12];2(3). Available from: http://www.cell.com/article/S2666675821000667/fulltext

111. Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.14.0 [Internet]. 2022. Available from: https://github.com/kevinblighe/EnhancedVolcano

112. Chen H, Boutros PC. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics [Internet]. BioMed Central; 2011 Jan 26 [cited 2022 Oct 12];12(1):1–7. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-35

113. Conesa A, Nueda MJ, Ferrer A, Talón M. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. Bioinformatics [Internet]. Oxford Academic; 2006 May 1 [cited 2022 Oct 12];22(9):1096–102. Available from:

https://academic.oup.com/bioinformatics/article/22/9/1096/200371

114. Walters MS, Gomi K, Ashbridge B, Moore MAS, Arbelaez V, Heldrich J, et al. Generation of a human airway epithelium derived basal cell line with multipotent differentiation capacity. Respir Res [Internet]. BioMed Central; 2013 Dec 3 [cited 2019 Feb 28];14(1):135. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24298994

115. Wang G, Lou HH, Salit J, Leopold PL, Driscoll S, Schymeinsky J, et al. Characterization of an immortalized human small airway basal stem/progenitor cell line with airway region-specific differentiation capacity. Respir Res [Internet]. BioMed Central Ltd.; 2019 Aug 23 [cited 2021 Jul 2];20(1):1–14. Available from: https://doi.org/10.1186/s12931-019-1140-9

116. McKimm-Breschkin JL. A simplified plaque assay for respiratory syncytial virus - Direct visualization of plaques without immunostaining. J Virol Methods. 2004;120(1):113–7.

117. KJ L, TD S. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods [Internet]. Methods; 2001 [cited 2021 Nov 3];25(4):402–8. Available from: https://pubmed.ncbi.nlm.nih.gov/11846609/

118. Essaidi-Laziosi M, Brito F, Benaoudia S, Royston L, Cagno V, Fernandes-Rocha M, et al. Propagation of respiratory viruses in human airway epithelia reveals persistent virus-specific signatures. J Allergy Clin Immunol [Internet]. Elsevier; 2018 Jun 1 [cited 2022 Oct 16];141(6):2074. Available from: /pmc/articles/PMC7112338/

119. Ilyushina NA, Ikizler MR, Kawaoka Y, Rudenko LG, Treanor JJ, Subbarao K, et al. Comparative Study of Influenza Virus Replication in MDCK Cells and in Primary Cells Derived from Adenoids and Airway Epithelium. J Virol [Internet]. American Society for Microbiology (ASM); 2012 Nov [cited 2023 May 26];86(21):11725. Available from: /pmc/articles/PMC3486302/

120. Cheng X, Zengel JR, Xu Q, Jin H. Surface glycoproteins of influenza A H3N2 virus modulate virus replication in the respiratory tract of ferrets. Virology. Academic Press; 2012 Oct 10;432(1):91–8.

121. Kim YM, Shin EC. Type I and III interferon responses in SARS-CoV-2 infection. Exp Mol Med 2021 535 [Internet]. Nature Publishing Group; 2021 May 6 [cited 2021 Dec 6];53(5):750–60. Available from: https://www.nature.com/articles/s12276-021-00592-0

122. Gonzàlez-Parra G, De Ridder F, Huntjens D, Roymans D, Ispas G, Dobrovolny HM. A comparison of RSV and influenza in vitro kinetic parameters reveals differences in infecting time. PLoS One [Internet]. Public Library of Science; 2018 Feb 1 [cited 2021 Nov 20];13(2). Available from: /pmc/articles/PMC5805318/

123. Talon J, Horvath CM, Polley R, Basler CF, Muster T, Palese P, et al. Activation of Interferon Regulatory Factor 3 Is Inhibited by the Influenza A Virus NS1 Protein. J Virol [Internet]. American Society for Microbiology (ASM); 2000 Sep [cited 2021 Nov 28];74(17):7989. Available from: /pmc/articles/PMC112330/

124. Sedeyn K, Schepens B, Saelens X. Respiratory syncytial virus nonstructural proteins 1 and 2: Exceptional disrupters of innate immune responses. PLOS Pathog [Internet]. Public Library of Science; 2019 [cited 2021 Dec 6];15(10):e1007984. Available from: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1007984

125. Dee K, Schultz V, Haney J, Bissett LA, Magill C, Murcia PR. Influenza A and respiratory syncytial virus trigger a cellular response that blocks severe acute respiratory syndrome virus 2 infection in the respiratory tract. J Infect Dis [Internet]. J Infect Dis; 2022 Dec 23 [cited 2023 May 16]; Available from: https://pubmed.ncbi.nlm.nih.gov/36550077/

126. Kim WM, Huang Y-H, Gandhi A, Blumberg RS. CEACAM1 structure and function in immunity and its therapeutic implications.

127. Kallal LE, Schaller MA, Lindell DM, Lira SA, Lukacs NW. CCR6 deficiency enhances immunity to respiratory syncytial virus by impairing recruitment of T helper type II-skewing dendritic cells. Eur J Immunol [Internet]. NIH Public Access; 2010 Apr [cited 2022 Oct 17];40(4):1042. Available from: /pmc/articles/PMC2952176/

128. Tregoning JS, Pribul PK, Pennycook AMJ, Hussell T, Wang B, Lukacs N, et al. The Chemokine MIP1α/CCL3 Determines Pathology in Primary RSV Infection by Regulating the Balance of T Cell Populations in the Murine Lung. PLoS One [Internet]. Public Library of Science; 2010 Feb 24 [cited 2022 Oct 18];5(2):e9381. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0009381

129. Gupta S, Leaf DE. Tocilizumab in COVID-19: some clarity amid controversy. Lancet [Internet]. Elsevier; 2021 May 1 [cited 2021 Dec 15];397(10285):1599–601. Available from: http://www.thelancet.com/article/S0140673621007121/fulltext

130. Yoshikawa T, Hill T, Li K, Peters CJ, Tseng C-TK. Severe Acute Respiratory Syndrome (SARS) Coronavirus-Induced Lung Epithelial Cytokines Exacerbate SARS Pathogenesis by Modulating Intrinsic Functions of Monocyte-Derived Macrophages and Dendritic Cells. J Virol [Internet]. American Society for Microbiology (ASM); 2009 Apr [cited 2021 Dec 5];83(7):3039. Available from: /pmc/articles/PMC2655569/

131. Mo LH, Luo XQ, Yang G, Liu JQ, Yang LT, Liu ZQ, et al. Epithelial cell-derived CD83 restores immune tolerance in the airway mucosa by inducing regulatory T-cell differentiation. Immunology [Internet]. Wiley-Blackwell; 2021 Jul 1 [cited 2022 Sep 4];163(3):310. Available from: /pmc/articles/PMC8207377/

132. Stanciu LA, Bellettato CM, Laza-Stanca V, Coyle AJ, Papi A, Johnston SL. Expression of programmed death-1 ligand (PD-L) 1, PD-L2, B7-H3, and inducible costimulator ligand on human respiratory tract epithelial cells and regulation by respiratory syncytial virus and type 1 and 2 cytokines. J Infect Dis. 2006 Feb 1;193(3):404–12.

133. Khairnar V, Duhan V, Patil AM, Zhou F, Bhat H, Thoens C, et al. CEACAM1 promotes CD8+ T cell responses and improves control of a chronic viral infection. Nat Commun [Internet]. Nature Publishing Group; 2018 Dec 1 [cited 2021 Dec 6];9(1). Available from: /pmc/articles/PMC6028648/

134. Persson BD, Jaffe AB, Fearns R, Danahay H. Respiratory Syncytial Virus Can Infect Basal Cells and Alter Human Airway Epithelial Differentiation. PLoS One [Internet]. 2014;9(7):e102368. Available from: http://dx.plos.org/10.1371/journal.pone.0102368

135. Hillyer P, Shepard R, Uehling M, Krenz M, Sheikh F, Thayer KR, et al. Differential Responses by Human Respiratory Epithelial Cell Lines to Respiratory Syncytial Virus Reflect Distinct Patterns of Infection Control. J Virol [Internet]. American Society for Microbiology (ASM); 2018 Aug [cited 2021 Aug 5];92(15). Available from: /pmc/articles/PMC6052282/

136. Rijsbergen LC, Lamers MM, Comvalius AD, Koutstaal RW, Schipper D, Duprex WP, et al. Human Respiratory Syncytial Virus Subgroup A and B Infections in Nasal, Bronchial, Small-Airway, and Organoid-Derived Respiratory Cultures. mSphere [Internet]. American Society for Microbiology (ASM); 2021 Jun 30 [cited 2021 Oct 10];6(3). Available from: /pmc/articles/PMC8125053/

137. Zhang L, Peeples ME, Boucher RC, Collins PL, Pickles RJ. Respiratory Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated Cells, and without Obvious

Cytopathology. J Virol [Internet]. American Society for Microbiology; 2002 Jun [cited 2021 Aug 3];76(11):5654–66. Available from: https://jvi.asm.org/content/76/11/5654

138. Sun Y, López CB. Preparation of Respiratory Syncytial Virus with High or Low Content of Defective Viral Particles and Their Purification from Viral Stocks. Bio-protocol [Internet]. NIH Public Access; 2016 May 20 [cited 2019 Jan 29];6(10). Available from: http://www.ncbi.nlm.nih.gov/pubmed/30079370