

University of Nottingham
UK | CHINA | MALAYSIA

Using Genetics and Cell Models to Understand Virus-host Interactions in the Airways

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2025

Christina Shrees

Student ID: 20312267

Supervisors: Professor Ian Sayers, Dr Nicholas Hannan, Dr Alexander Tarr

School of Medicine

University of Nottingham

Table of Contents

Abstract 1

Acknowledgements 2

Abbreviations 3

1.0 Introduction 6

1.1 The Respiratory System 7

1.1.1 The Bronchiolar Airway Epithelium 9

1.2 Respiratory Viruses 13

1.2.1 Rhinovirus 21

1.2.2 Respiratory Syncytial Virus 28

1.2.3 SARS-CoV-2 36

1.3 Host Genetic Factors Associated with Respiratory Viruses 44

1.3.1 Rhinovirus 45

1.3.2 RSV 47

1.3.3 SARS-CoV-2 64

1.3.4 Summary 76

1.4 Modelling the Airway Epithelium to Study Host-Virus Interactions 78

1.4.1 Primary Cell Culture 78

1.4.2 Continuous Cell Lines 79

1.4.3 Induced Pluripotent Stem Cells 80

1.5 Using Transcriptomics to Understand Host-Virus Interactions 82

1.6 Aims and Objectives 84

1.6.1 Overall Aim 84

1.6.2 Objectives 84

2.0 Materials and Methods 85

2.1 Molecular Biology Methods 85

2.1.1 Material and Reagents 85

2.1.2 RNA Extraction 87

2.1.3 Quantification of RNA 88

2.2 Cell Culture Methods 89
2.2.1 Materials and Reagents..... 89
2.2.2 HBECs..... 93
2.2.3 HEK293-T..... 97
2.3 Viral Methods..... 98
2.3.1 Propagation of RSV-A16 and RV-A16 98
2.3.2 Determination of Multiplicity of infection (MOI) for Respiratory Syncytial Virus and Rhinovirus 99
2.3.3 Infection of iPSCs and HBECs with RSV-A2 and RV-A16 99
3.0. Translating Host-respiratory Virus Genetic Susceptibility and Severity Signals to Candidate Casual Genes and Pathways..... 100
3.1 Introduction 100
3.2 Aims and Objectives 105
3.3 Methods..... 106
3.3.1 Literature Search of SNPs Associated with Host-Virus Interactions 107
3.3.2 Analysis of Signal-to-Gene Association 109
3.3.3 Analysis of PheWAS Using GeneATLAS UK Biobank Data..... 111
3.3.4 Analysis of Total and Single Cell RNA Expression in Lung Tissue 111
3.3.5 Analysis Using Ingenuity Pathway Analysis (IPA)..... 112
3.3.6 Analysis of Drug Gene interaction 112
3.4 Results..... 113
3.4.1 Analysis of Signal-to-Gene Association Identifies 6, 108 and 119 Candidate Causal Genes Associated with Host-RV/ RSV/ SARS-CoV-2 Interactions Respectively 113
3.4.2 Cell Type Specificity Was Observed Through Analysis of Single Cell RNA Expression in Lung Tissue 120
3.4.3 Signals Show Association with Respiratory, Blood Cell Counts, Allergy, Inflammatory and Auto-immune Disorders..... 127
3.4.4 IPA Analysis Identifies Pathways Associated with Host-Virus Interactions..... 133

3.4.5 IPA Analysis Identifies Diseases and Functions Associated with Host-Virus Interactions 139
3.4.6 Upstream Regulators of Genes Identified for Host-RV/RSV/SARS-CoV-2 Interactions 141
3.4.7 Analysis of Gene-Drug Interaction and Clinical Trial Data Identifies Genes as Druggable Targets 143
3.5 Discussion 144
4.0 Using Transcriptomics to Study How Primary Human Epithelial Cells Respond to Prolonged Virus Infections 152
4.1 Introduction 152
4.2 Aims and Objectives 156
4.3 Methods and Materials 157
4.3.1 Viral Infection and RNA isolation 157
4.3.2 RNA-Seq Methods 158
4.4 Results 161
4.4.1 How Do Primary Human Epithelial Cells Respond To RSV-A2 and RVA-16 In Prolonged Infections? 161
4.4.2 What Are the Similarities and Differences in the Way Human Epithelial Cells Respond To Prolonged RSV-A2 and RVA-16 Infections? 180
4.4.3 Does The Viral Gene Signature Provide Support For Genes Supported By Genetics? 199
4.5 Discussion 204
5.0 Development of Induced Pluripotent Stem Cells and Human Bronchial Epithelial Cells-Derived Models of the Airway Epithelium to Investigate Host-Virus Interactions 212
5.1 Introduction 212
5.1.1 Induced Pluripotent Stem Cells-Derived Airway Epithelium 213
5.1.2 Primary Cells-Derived Airway Epithelium 222
5.1.3 Studying Host-Virus Interactions Using Cell Culture Models 223
5.2 Aims and Objectives 225

5.3 Methods and Materials 226
5.3.1 Molecular Biology Methods 226
5.3.2 Cell Culture Methods 231
5.3.3 Viral Methods 247
5.3.4 Immunohistochemistry Methods 248
5.3.5 Proteomic Methods 251
5.3.6 Statistical Analysis 251
5.4 Results 252
5.4.1 LOPCK and REBL-PAT iPSCs Were Differentiated Into Stage 4 Bronchial Lung Progenitor Cells 252
5.4.2 D195307 HBECs Were Differentiated Into Mucociliary Pseudostratified Epithelium 259
5.4.3 LOPCK and REBL-PAT iPSCs Were Unsuccessful in Forming a Mucociliary Pseudostratified Epithelium 265
5.4.4 Characterisation of LOPCK and REBL-PAT iPSCs at Different Stages of Differentiation 266
5.4.5 Investigating How Permissive iPSCs and HBECs Are To RSV-A2 At Various Stages Of Differentiation 274
5.4.6 Comparison Between The Way Human Epithelial Cells Respond To RV, RSV and SARS-CoV-2 Infections in ALI-HBECs Model of the Airway Epithelium 286
5.5 Discussion 294
6.0 General Discussion 301
6.1 Aims and Objectives 301
6.2 Main Findings 303
6.3 Study Limitations 308
6.4 Future Work 310
6.5 Final Conclusions 312
7.0 Bibliography 314

Abstract

Respiratory virus infections are a major cause of mortality and morbidity worldwide. It is important to investigate the genetic and molecular background underlying host-virus interactions. Identifying genes and pathways that may modify viral pathogenesis could facilitate the development of therapeutics to treat viral infections. Three approaches were used in the thesis to understand host-virus interactions: genomics, transcriptomics and cell models. Signal-to-gene association analysis was conducted to translate genetic signals associated with susceptibility or severity to rhinovirus (RV), respiratory syncytial virus (RSV), and SARS-CoV-2 infection. 6, 108 and 119 candidate causal genes were identified for RV, RSV and SARS-CoV-2 respectively. IL6-, IL-AS1 and STEAP1B was found to overlap between the three viruses. Downstream analysis of the genes revealed that immunomodulatory pathways such as the Th1 signalling pathway may be particularly important for the three viruses. Drugs currently in Phase I, II and III of clinical trials which target the candidate causal genes were identified. Transcriptomics was utilised to investigate how human bronchial epithelial cells (HBECs) respond to prolonged RSV-A2 and RV-A16 infection. Cell cycle pathways were predicted to be activated upon prolonged RSV infection whereas inflammatory pathways were predicted to be downregulated. Keratinisation was the top pathway predicted to be activated whereas elastic fibre formation pathway was predicted to be downregulated in prolonged RV-A16 infection. To further understand host-virus interaction, induced pluripotent stem cell (iPSC) and HBEC air-liquid interface models of the airway epithelium were used to study RSV infection. Overall, these findings provide insight into genes and pathways important for host-virus susceptibility, and opportunities for development of therapeutics.

Acknowledgements

Firstly, I would like to thank my supervisors Professor Ian Sayers, Dr Nicholas Hannan and Dr Alexander Tarr for giving me this opportunity to undertake the PhD and for their continual guidance, support and advice during the PhD process.

I would like to thank Michael Portelli, Robert Hall and Kamini Rakkar for their kindness and patience when helping me train in the techniques and for their advice and support throughout the PhD. I would like to thank everyone in Ian Sayers lab group, Nicholas Hannan Lab Group and Alexander Tarr Lab group for their support and guidance. I would also like to thank everyone in Respiratory Medicine and the office for their friendship and support.

I would like to thank my parents, brother, friends and family for their continual support and encouragement. I really appreciate all that you guys have done for me as the PhD was not an easy journey for me. I would also like to thank my friends for their emotional support and for always being there for me.

Abbreviations

AFE	Anterior Foregut Epithelium
ANOVA	Analysis of Variance
AOM	Acute Otitis Media
ARS	Acute Rhinosinusitis
cDNA	Complementary DNA
CDHR3	Cadherin-related Family Member 3
COPD	Chronic Obstructive Pulmonary Disease
COVID-19	Coronavirus Disease 2019
CT	Cycle Threshold
DNA	Deoxyribose Nucleic Acid
DE	Definitive Endoderm
eQTL	Expression Quantitative Trait Loci
FOXA2	Forkhead Box Protein A2
FOXJ1	Forkhead Box Protein J1
GWAS	Genome Wide Association Studies
HAdV	Human Adenovirus
HBEC	Human Bronchial Epithelial Cells
HBoV	Human Bocavirus
HMPV	Human Metapneumovirus
HPIV	Human Parainfluenza Viruses
HSPG	Heparan Sulphate Proteoglycans
ICAM-1	Intracellular Adhesion Molecule-1
IFN	Interferon

IL	Interleukin
KRT5	Keratin 5
LD	Linkage Disequilibrium
LRI	Lower Respiratory Infection
MUC5AC	Mucin 5AC
NF-κB	Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells
NKX2.1	NK2 homeobox 1
ORF	Open Reading Frames
PCR	Polymerase Chain Reaction
PNEC	Pulmonary Neuroendocrine Cells
qPCR	Quantitative PCR
RNA	Ribonucleic Acid
RNA-Seq	RNA-Sequencing
RSV	Respiratory Syncytial Virus
RT-qPCR	Reverse Transcription Quantitative PCR
RVI	Respiratory Virus-induced Infection
RV	Rhinovirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SOX17	SRY-box 17
SNP	Single Nucleotide Polymorphism
TLR	Toll-like Receptors
URI	Upper Respiratory Infection
URTI	Upper Respiratory Tract Infection
VPG	Viral Priming Protein

WHO	World Health Organisation

1.0 Introduction

Infections attributed to respiratory viruses are a major cause for mortality and morbidity worldwide. It has a lasting impact on healthcare, education, economy, and wellbeing of individuals suffering from infection. Therefore, it is vital to understand the molecular, genetic and transcriptomic basis underpinning host-virus interactions. This can facilitate identification of genes and pathways important during pathogenesis. The similarities and differences between respiratory viruses can also be examined by exploring their genetics, transcriptomics and cell biology via cell culture models of the bronchiolar airway epithelium. Furthermore, potential therapeutic strategies which can be applied across viruses can be developed. As new strains of viruses are constantly being discovered, sustained development of preventatives and anti-viral therapeutics is vital. Utilising -omics based approach for drug discovery will drive innovation towards therapeutics to target respiratory viruses and enhance preparation for future pandemics.

1.1 The Respiratory System

The primary role of the respiratory system is to facilitate gaseous exchange, but it also has critical function in the maintenance of the respiratory homeostasis[1]. To maintain this homeostasis, the airways perform various barrier functions[2]. The airways are lined with a mucosal layer which acts as a moisture barrier[2]. To limit excessive moisture loss during ventilation, the mucosal lining helps increase the air humidity in the upper airway[2]. Temperature barrier is exerted and is dependent on the body temperature. A rich lymphatic system lines the airways and acts as a barrier to infection[2]. This lymphatic system includes the mucosa-associated lymphoid tissue[2]. Moreover, macrophages circuit around the respiratory surfaces as an essential part of the air-blood barrier. The respiratory system is divided into two tracts: the upper respiratory tract consisting of the nose, nasal cavity, pharynx and the larynx, and the lower respiratory tract composed of the trachea, bronchi, bronchioles, alveolar duct, and alveoli[3]. An important role of the nasal cavity is to induce turbulent airflow of inspired air[4]. This helps infectious particles and debris to be deposited and entrapped before reaching the lung[4]. As a result, primary infection of respiratory viruses is typically established in the nasal epithelium[4]. Therefore, they serve as a dynamic, front-line defensive barrier[4]. Maximal deposition density is achieved by aerosolised viral particles in the nasal cavity where viruses initially replicate[5]. The viruses then spread to the lower airways where aspiration is mediated along a nasal/oral-lung aspiration axis[4]. Dissemination of the viruses to the lower airway is limited if an efficient innate immune response is induced[4].

The nasal and oral cavity is connected to the larynx and the oesophagus by a tube-like passage called the pharynx[6]. Located between the pharynx and the trachea, the larynx is an organ involved in phonation, breathing and protection of the trachea[7]. The trachea, which is composed of connective tissue, smooth muscle, cartilage, and mucosal layer, connects the larynx to the principal bronchi[8]. Some of the main functions of the trachea include air flow into the lungs, humidification of air and mucociliary clearance[8]. As an extension of the trachea, the bronchi serve as the main passageway into the lungs[8]. The trachea bifurcates into the bronchi and together they make up the

tracheobronchial tree of the lungs[8]. The airways of the bronchi branch out into three different types of bronchioles: conducting, terminal and respiratory with respiratory bronchioles giving rise to alveoli[8]. Alveoli are lined with a layer of surfactant producing pneumocyte[2]. Alveolar sacs bud from alveolar ducts, and they form clusters[2]. These alveoli clusters are connected by pores to allow equalisation of air pressure[2]. They form the air-blood barrier in conjunction with capillaries[2].

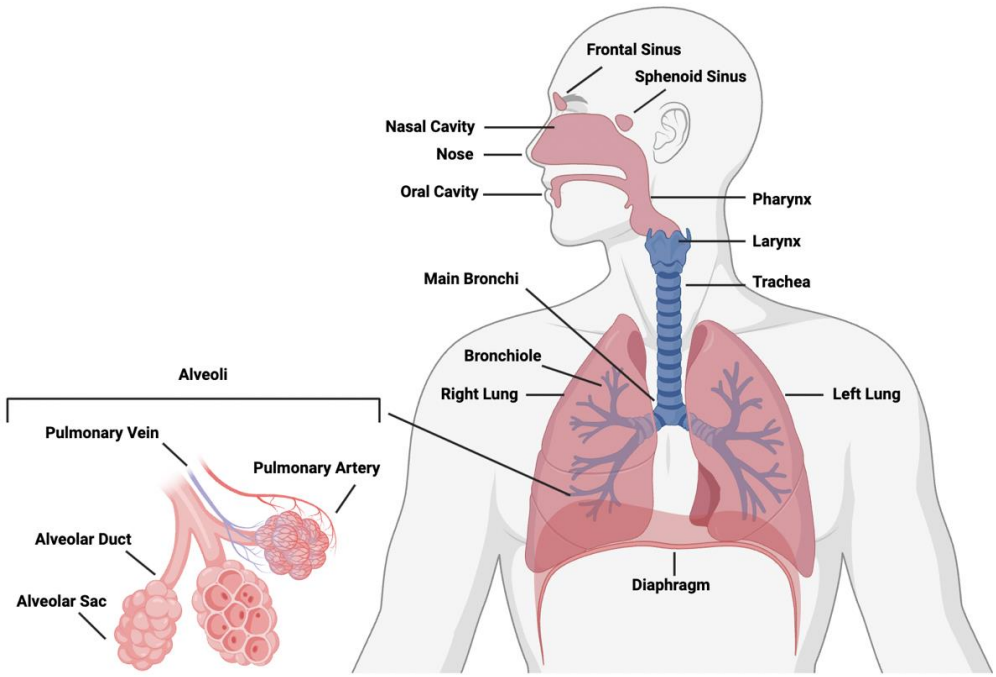


Figure 1-1. The Respiratory System. The upper respiratory system is composed of the nose, nasal cavity, sinus, pharynx, larynx. The trachea, lung, bronchi, bronchioles, and alveoli make up the lower respiratory system. Diagram was created using BioRender with Ball *et al* as reference.

1.1.1 The Bronchiolar Airway Epithelium

Lining the luminal surface of the tracheobronchial tree, the bronchial airway epithelium acts as a barrier to protect the lung stressors such as pathogens, toxic substances, and irritants[9]. Pseudostratified epithelium consisting of different specialised cells lines the proximal airway[10]. A physical barrier is maintained by intercellular adhesion complex whereas the tight junction proteins zonula occludens found in the apices of adjacent cells constrain diffusion of electrolytes and molecules[11]. The structural integrity of the airway epithelium is supported by desmosomes and gap junctions[11]. The bronchiolar airway epithelium also provides a plentiful source of arachidonic acid metabolites which aids the regulation of mucus secretion, smooth muscle tone, and release of neurotransmitters[11]. Multiple different epithelial cell types of the bronchiolar airway epithelium work in conjunction with the immune system to prevent stressors from infiltrating the submucosal compartment[12].

1.1.1.1 Basal Cells

Basal cells present a cuboidal morphology and are distinguished by the expression of cytokeratin 5 and 14, and transformation-related protein 63 (Trp63)[1]. They are attached to basement membrane via hemidesmosomes and they in turn promote the attachment of other superficial cells through hemidesmosomal complexes[11]. An essential component for the regeneration of the airway epithelium, basal cells remain relatively quiescent but become activated in response to injury[9]. They act as progenitor cells and are to differentiate into different epithelial cell types to sustain homeostasis following injury[12]. Basal cells then acquire damage-associated phenotypes required for the regeneration along with matrix metalloproteinases, mesenchymal cell-associated vimentin, and cytokeratin 6, 13 and 14 for the formation of the barrier[9]. The restoration of the epithelium is facilitated by the ability of the basal cells to differentiate into specialised cell type such as ciliated, club and goblet cells[1].

1.1.1.2 Ciliated Cells

As the predominant cell type, ciliated cells make up 50% of all epithelial cells in the airways[11]. Ciliated cells possess hair-like projections on the apical surface called cilia. Cilia are composed of a microtubule scaffold called the axoneme where a central pair of singlet protofilaments is surrounded by a ring of nine doublet microtubule giving a '9 + 2+ structure[13]. The microtubule pairs own dynein arms that enable motility through ATP hydrolysis[13]. This facilitates mucociliary clearance where pathogens and debris are trapped and expelled to maintain the airway homeostasis[1]. Differentiation into ciliated cells is dictated by the absence of notch signalling where geminin coiled-coil domain-containing Protein 1 (GEMC1) induces the expression of multiciliate differentiation and DNA synthesis associated cell cycle protein (MCIDAS)[13]. This then activated the expression of the transcription factor forkhead box protein J1 (FOXJ1) [1]. FOXJ1 helps regulate the docking of basal bodies to the apical surface, cilia elongation and synthesis of motile components[13].

1.1.1.3 Goblet Cells

Goblet cells are the primary mucus producing cells of the airway epithelium. The heavily glycosylated mucus consists of membrane-bound acidic-mucin granules which are secreted to trap foreign particles in the lumen of the airway[11]. Three classes of mucins exist in the airways: gel-forming mucins 5AC (MUC5AC) and MUC5B, cell surface associated mucins MUC1, MUC4, MUC16 and MUC20, and secreted mucins that do not polymerise such as MUC7[14]. MUC5AC is predominantly expressed in the goblet cells whilst MUC5B is expressed in the submucosal glands[12]. The goblet cells mediate constitutive mucus secretion[12]. Whereas submucosal glands, which are controlled by autonomic innervation, can increase mucus production rapidly in response to stimulation[12]. The mucins can undergo modification at the epigenetic, transcriptional, post-transcription, and post-translational level to enable binding of mucins and clearance of pathogens[14]. Differentiation into goblet cells and production of mucus are activated by sam-pointed domain-containing Ets-like factor (SPDEF) and forkhead box A3 (FOXA3)[1].

1.1.1.4 Club Cells

Club cells have an immunomodulatory role, and they contain membrane-bound electron dense secretory granules[15]. These granules secrete club cell secretory protein (CCSP) which has an anti-inflammatory and anti-protease activity[16]. Club cells can also secrete surfactant proteins and anti-proteinases such as secretory leukocyte protease inhibitor to help protect the integrity of the bronchiolar epithelium[15]. In addition, club cells produce p450 mono-oxygenase which metabolises xenobiotic compounds such as aromatic hydrocarbons found in cigarette smoke[17]. Furthermore, club cells serve as primary progenitors of ciliated cells[13].

1.1.1.5 Brush Cells

Brush cells are chemosensory epithelial cells which coordinate interaction with the external environment[18]. They share a close interaction with nerve fibres to mediate communication between the neuronal and immune pathways[19]. The cells have a pear/bottle shape with a wide base and a narrow apex with a tuft of blunt, squat apical microvilli which gives them a distinctive morphology[20]. The microvilli possess filaments which stretch into the underlying cytoplasm[20]. The function of brush cells is not fully understood[21]. However, it has been suggested that they have a role in virulence associated clearance and promoting mucociliary clearance by releasing acetylcholine[21].

1.1.1.6 Pulmonary Neuroendocrine Cells

Pulmonary neuroendocrine cells (PNECs) are the first cell type to form and differentiate within the airway epithelium during lung development[22]. During the neonatal phase, the number of PNECs are at its peak[22]. The PNEC system represents the lung stem cell niche which is key to the regeneration of the airway epithelium during the postnasal phase[23]. PNECs are rare cells and there is approximately one PNEC per 2500 epithelial cells in a healthy individual[17]. PNECs have a pyramidal shape and are typically tall[17]. They also possess an apical microvilli and extend from the basal lamina of the epithelium[17]. Neural elements and smooth muscles interact with PNECs to mediate the regulatory properties and the bronchomotor and to maintain vascular tone[24]. In addition, PNECs secrete biogenic amines and peptides that

are important in airway function and foetal lung growth[23]. It has also been suggested that PNEC may act as hypoxia sensitive airway chemoreceptors and play a role in mediating localised epithelial cell growth and regeneration[22].

1.1.1.7 Ionocytes

Pulmonary ionocytes account for 1-2% of human airway epithelial cells and are more enriched in cystic fibrosis transmembrane conductance regulator (CFTR) messenger ribonucleic acid (mRNA) than in any other airway cell type[18]. Through the combination of scRNA-seq and 'pulse-seq' where in vivo lineage was traced over three different time points, basal cells were found to be the primary source of ionocytes[25]. Moreover, further investigation of rare cell lineages using CRISPR-Cas9 knockout system showed that ionocytes and PNECs arise from POU domain class 2 transcription factor 3 $(POU2F3)^{+}$ tuft-like cells during differentiation[26]. Electrophysiological studies have shown positive correlation between the number of ionocytes and CFTR-mediated CL^{-} secretion in human bronchial epithelial cells (HBECs)[27]. Ionocytes have shown to regulate airway surface liquid (ASL) in HBECs[28]. Furthermore, abnormal function of ionocytes and club cells contributes to dehydrated and acidic ASL in individuals with cystic fibrosis[28]. Despite their rarity, this shows the importance of ionocytes in mediating transepithelial Cl^{-} secretion[27].

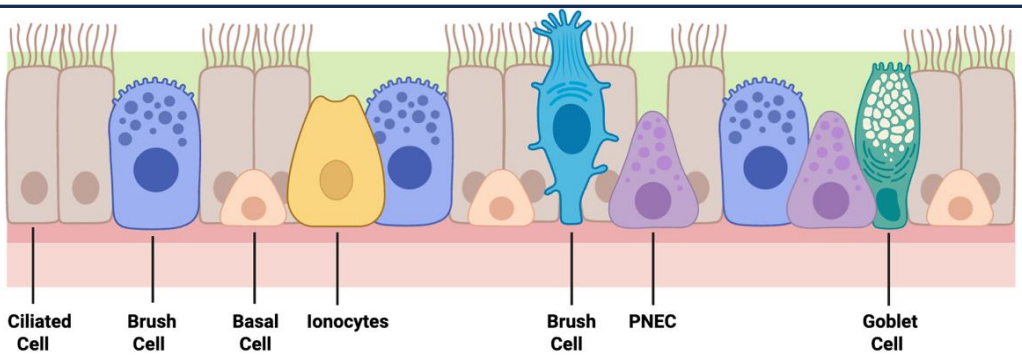


Figure 1-2. The Airway Epithelium. Schematic of the bronchiolar airway epithelium. The airway epithelium is a pseudostratified layer that has been differentiated to produce various specialised cells such as basal, ciliated, goblet, club, brush cells, PNEC and ionocytes. Diagram was created using BioRender with Davis *et al* as reference.

1.2 Respiratory Viruses

Respiratory virus-induced infections (RVIs) are one of the major causes of mortality and morbidity globally and causes significant social and economic effect[29]. Rhinovirus (RV), respiratory syncytial virus (RSV), coronaviruses and influenza are some of the most common and important respiratory viruses[30]. In 2015, lower respiratory infections (LRIs) caused by respiratory viruses were estimated to be responsible for 2.74 million deaths worldwide[31]. Furthermore, 704,000 deaths in children less than 5 years of age were attributed to LRIs[31]. LRIs resulted in 60.6 million disability-adjusted life-years globally[31]. However, only a few therapeutics against respiratory infections exist, therefore understanding the mechanisms behind pathogenesis and interaction between host and virus is important for respiratory disease modelling and the development of novel therapeutics. This section summarises an overview of the following respiratory viruses: human adenovirus (HAdVs), human bocavirus (HBoV), human metapneumovirus (HMPV) and human parainfluenza viruses (HPIV). A focused review on rhinovirus (RV), respiratory syncytial virus (RSV) and severe acute respiratory syndrome-corona virus 2 (SARS-CoV-2) is detailed in Chapter 1.2.1.

HAdVs comprises a large group of double-stranded deoxyribonucleic acid (dsDNA) viruses belonging to the *Adenoviridae* family[32]. HAdV are classified into 51 distinct serotypes and distributed across six species: A-F The virions are around 90 nm in size, non-enveloped with an icosahedral symmetry[33]. The capsid consists of three types of capsomere proteins : hexons, penton bases, and pento fibres[34]. The genome is linear, non-segmented dsDNA between 26-46 kb pairs[33]. Most human HAdVs gain entry by using the coxsackie B and adenovirus receptor whereas HAdV species B used CD46[33]. During infection, HAdVs bind to a cellular receptor via the penton fibre. This interaction triggers endocytosis and exposes the virus to the cytosol. HAdV then associates with the microtubule network in the cytosol and is trafficked to the nucleus via the dynein motors. The final disassembly occurs at the nuclear pore where the remaining capsid proteins interact with the nuclear pore complex to deposit dsDNA genome into the nucleus[33]. A set of ‘early’ mRNAs are transcribed in the nucleus[34]. These mRNAs direct translation of proteins that facilitate cell

cycle and DNA replication and impedes apoptosis. Late mRNAs are transcribed to produce new viral proteins[34]. Viral particles are assembled in the nucleus and released through cell lysis.

A wide spectrum of diseases is caused by HAdV infection however respiratory infections remain the most common especially in children under 5 years. Around 5-7% of paediatric respiratory tract infections were estimated to be attributed to HAdV globally in 2019[33]. In 2009 however infections by serotypes 1-3 and 5 accounted for 5-20% of URI and 5% of LRI cases in children[34]. Serotypes 1-7 and 21 are most frequently associated with respiratory diseases[34]. Infections by serotypes 1-3 and 5 accounted for 5-20% of URI and 5% of LRI cases in children in 2009[34]. HAdVs can also cause infections of the gastrointestinal tract and conjunctiva[32]. Although there is no clear seasonality in terms of HAdV infection, most epidemics tend to occur in winter or early spring[32]. Outbreaks often occur in small populations such as hospitals, and especially in military installations[32]. Serotypes 4 and 7 account for 6-16% attack rate per week in newly assembled military recruits[34]. Overall attack rates may reach up to 80% with 20-40% of individuals requiring hospitalisation[34]. During the summer in temperate climates, pharyngoconjunctival fever caused by serotypes 3 and 7 can be epidemic or endemic among children. Moreover, HAdV infection can result in epidemics of febrile respiratory illness, gastrointestinal and diarrheal illness, and keratoconjunctivitis[32]. Severe and disseminated HAdV infection rarely occurs in immunocompetent hosts but can occur in immunocompromised patients[32]. The primary management strategies for HAdV infection revolve around supportive care and reduction of immunosuppression when relevant[35]. Currently, there is no specific routine anti-viral treatment for HAdV infection[35]. Broad-spectrum anti-virals such as ganciclovir, cidofovir, and ribavirin are instead used to treat severe infection[35].

HBoV is an autonomously replicating human parvovirus and belongs to the family *Parvoviridae*[36]. They are small single-stranded DNA viruses with a genome of 5-5.5 kb and a particle size from 18-26 nm[33]. The genome consists of three open reading frames (ORFs) with ORF1 and ORF2 encoding four non-structural proteins (NS1-4) and NP1 respectively[33]. The capsid proteins VP1-3 are encoded by ORF3. The virions are non-enveloped and composed of up

to 60 capsid proteins[33]. Clinical manifestation of infection includes rhinitis, bronchiolitis, bronchitis, asthma exacerbation, and pneumonia[34]. BoV infection can range from mild to life-threatening in children and immunocompromised individuals[36]. The distribution of HBoV infection is global with cases occurring throughout the year particularly during winter and spring[37]. Moreover, a high degree of co-infection with other respiratory viruses have been detected in the cases[33]. Children between 6-24 months suffer from the greatest disease burden with HBoV being detected in up to 20% of acute respiratory infection[33]. HBoV infections have shown to be highly disruptive. They perturb tight-junction barriers of the airway epithelium resulting in the epithelium thinning, loss of cilia and hypertrophy of the cells[38]. It has been suggested that HBoV elicits an inflammatory response leading to pathogenesis and cell death via apoptotic and pyroptotic mechanisms[33]. There is no specific treatment for HBoV infection[34]. However, treatment with high-dose glucocorticoid has shown to result in rapid recovery from severe pulmonary disease caused by HBoV[39].

Influenza viruses belong to the *Orthomyxoviridae* family of RNA viruses[40]. The virus is enveloped with a genome consisting of a segmented negative-strand ssRNA. The RNA encodes for hemagglutinin (HA) and neuraminidase (NA) surface glycoproteins. HA facilitates binding of the virus to its α2,6-linked sialyloligosaccharide receptors on host cells. Whereas NA mediates release of viral particles by cleaving sialyloligosaccharide residues from the surface of host cells. There are four categories of influenza virus strains: A, B, C and D. Influenza A causes mild to severe illness and is the most common strain during seasonal epidemics. Whereas influenza B causes highly contagious infection which can develop into serious illness. Furthermore, type B has been responsible for minor localised outbreaks[40]. Localised outbreaks are clusters of infections that are confined to settings such as hospitals and care homes[40]. Although influenza C primarily infects humans, the virus can also infect animals including pigs[41]. Type C causes mild upper respiratory symptoms, minor localised outbreaks and sporadic cases. In contrast influenza D is not known to cause infection in humans. However, type D affects cattle and pigs. Type A is classified into subtypes on the basis of antigenic variation in HA and NA antigens[42]. Sixteen antigenic variants of HA (H1-H16) and nine variants of

NA (N1-N9) have been identified[42]. Both type A and B accumulate point mutations in HA and NA genes due to limitations in the proofreading ability of viral RNA-dependent RNA polymerase. This along with antibody-mediated selection pressure results in antigenic drift which allows the virus to escape immunity elicited by vaccination or previous exposures. This gives rise to seasonal epidemics[43].

Globally, around 4 million severe infections and half a million deaths are attributed to influenza according to the World Health Organisation (WHO)[40]. Infection occurs in individuals of any age group[40]. However, people over the age of 65 are at higher risk of hospital admission, risk of severe complications, and death[44]. Complications such as pneumonia can arise from influenza during pregnancy[45]. This is associated with higher risk of hospital admission and disabilities in mothers[45]. Currently available vaccines include live attenuated influenza vaccine, inactivated influenza vaccine, and recombinant influenza vaccine. Up to 60% efficacy of vaccines against type A and B infection have been reported by some studies. Supportive treatment such as hydration and anti-inflammatory drugs is offered in uncomplicated infection and low-risk cases. Currently four anti-viral are approved for preventing and treating infection: oseltamivir (Tamiflu), baloxavir marboxil (Xofluza), peramivir (Rapivab), and zanamivir (Relenza). Oseltamivir, peramivir and zanamivir are NA inhibitors and block the function of NA to prevent the virus from leaving infected cells. Whereas baloxavir marboxil is an inhibitor of cap-dependent endonuclease[46]. The drug impedes transcription of viral RNA and prevents the virus from replicating[46].

HMPV is a negative sense ssRNA belonging to the subfamily *Pneumoviridae* within the family *Paramyxoviridae*[47]. HMPV particles are approximately 200 nm in diameter and infect nasal and lung epithelial cells[33]. The particles are enveloped and pleomorphic[34]. During HMPV infection, F fusion glycoprotein binds to heparan sulphate and other glycosylated proteins. The pH lowers resulting in the activation of the fusion protein[33]. This promotes fusion between the viral and cellular membrane. The genetic material is released into the cytosol and viral proteins are produced[33]. HMPV circulates mainly in late winter and spring in temperate climates. HMPV infections can range from mild upper

acute respiratory infection (ARI) to pneumonia and severe bronchiolitis. Globally, approximately 20% of Respiratory Tract Infection (RTIs) in children are associated with HMPV[33]. The peak age of hospitalisation is 6-12 months[34]. Furthermore, more than 90% of children have antibodies to HMPV by the age of 5[34]. Whereas the seroprevalence is around 100% among adults[34]. Mild upper respiratory tract infection (URTI) has been reported in healthy children and adults with HMPV infection. However severe symptoms can develop in immunocompromised adults and the elderly[33]. Pathologic studies have shown sloughing of epithelial cells with eosinophilic cytoplasmic inclusions, formation of hyaline membranes, multinucleated giant cells, diffuse alveolar damage, and acute and organising lung injury[34]. Supportive measures such as corticosteroids, oxygen therapy, mechanical ventilation and bronchodilators are utilised to manage and control HMPV infection. However, no specific anti-viral treatment exists for HMPV.

HPIVs are genetically and antigenically divided into type 1-4. HPIV types 1 and 3 are classified into the genus *Respirovirus* whilst types 2 and 4 are grouped into *Rubulavirus*. Types 1-3 are a major cause of LRIs in infants, young children, immunocompromised and chronically ill individuals, and the elderly[48]. HPIV are non-segmented, negative sense ssRNA virus[48]. The genome is approximately 15 nm, and the virion is enveloped and pleomorphic, ranging from 150-200 nm in diameter[33]. The hemagglutinin-neuraminidase glycoprotein with HA and NA activities is located in the virion envelope[48]. HA is attached to the sialic acids on the surface of permissive cells. The surface proteases cleave the fusion protein which induces a conformational change which stimulates the fusion between the viral envelope and the plasma membrane[33]. Following the fusion reaction, the viral genome is released into the cytoplasm where replication occurs using a virus-specific RNA-dependent RNA polymerase[48]. Viral mRNAs are translated into viral proteins through the cellular ribosomal machinery[48].

HPIV infection is responsible for one-third of RTIs in the United States[33]. Types 1 and 3 are more commonly associated with LRIs in children, immunocompromised and chronically ill individuals, and the elderly[34]. Whereas type 4 mainly causes upper respiratory infections (URIs) in children

and adults[34]. Serological studies have reported positive seroprevalence for HPIV by the age of 6-10 years for most children[33]. The symptoms can range from mild common cold and fever to more severe disease such as bronchiolitis, croup and pneumonia[33]. Croup is the most common cause of HPIV-induced hospitalisation whereas bronchiolitis is also common in particular during the first year of life[33]. HPIV-3 is the second most common cause for bronchiolitis and pneumonia in young children after RSV[33]. Furthermore, young, immunocompromised children are at the greatest risk for developing severe HPIV infection and can ultimately be fatal[33]. Immunocompromised individuals and elderly are at greatest risk for HPIV re-infection[33]. There is no anti-viral agent currently on the market with proven efficacy for HPIV. Corticosteroids are prescribed to treat to croup and nebulisers to alleviate respiratory symptoms[48]

Table 1-1. Table of the Main Respiratory Viruses. The table outlines the family, genome size and type, structure of the virion, and major clinical disease caused by the viruses.

Virus	Family	Genome Type	Genome Size	Structure of Virion	Major Clinical Disease
Human Adenovirus (HAdV)	*Adenoviridae*	dsDNA	35 kb	90 nm, Non-enveloped	URTI, Gastroenteritis, Conjunctivitis, Myocarditis
Human Bocavirus (HBoV)	*Parvoviridae*	ssDNA	5-5.5 kb	18-26 nm, Non-enveloped	LRTI
Human metapneumovirus (HMPV)	*Paramyxoviridae*	-ssRNA	13 kb	150–200 nm, Enveloped	URT symptoms
Human parainfluenza viruses (HPIV)	*Paramyxoviridae*	-ssRNA	15 kb	150–200 nm, Enveloped	Bronchiolitis, Croup

Virus	Family	Genome Type	Genome Size	Structure of Virion	Major Clinical Disease
Influenza A (IAV)	*Orthomyxoviridae*	-ssRNA	13.5 kb	80-120 nm, Enveloped	Influenza
Rhinovirus (RV)	*Picornaviridae*	+ssRNA	7.2 kb	30 nm, Non-enveloped	Cold, Asthma Exacerbation
Respiratory Syncytial Virus (RSV)	*Pneumoviridae*	-ssRNA	15.2 kb	120-200 nm, Enveloped	Bronchitis, Wheezing, Asthma Exacerbation
SARS-CoV	*Coronaviridae*	+ssRNA	27-32 kb	120 nm, Enveloped	LRTI, ARDS

1.2.1 Rhinovirus

RVs are single-stranded RNA (ssRNA), positive-sense viruses. RV belongs to the *Enterovirus* genus of the *Picornaviridae* family[33]. As the most commonly isolated respiratory virus from individuals with mild UTRI, RV is responsible for between a half and two-thirds of colds[33]. RVs are a frequent causative agent for upper respiratory tract infection, sinusitis, and otitis media[49]. RV has also been implicated in a range of different lower respiratory tract infections such as driving exacerbation of asthma and chronic obstructive pulmonary disease (COPD) in adults, wheezing in children and pneumonia[49]. Respiratory virus accounts for approximately 80% of the asthma exacerbation with RV attributing to two thirds of the cases. RV-induced wheezing in early age is shown to be a predictor for the development of asthma[50]. It is theorised that stimulation of airway inflammation and modelling by RV facilitates development of asthma[50]. Furthermore, medical visits and absent days in school and work attributed to RV costs billions of dollars annually in the United States causing a substantial effect on the economy[49]. Currently, no approved antivirals or vaccinations are available, and the treatments are typically palliative[51].

1.2.1.1 Structure and Genotypes/Serotypes

The RV genome is a monocistronic uncapped ssRNA with a size of 7-8 kb[49]. A single open reading frame (ORF) is located in the genome which is joined by a 5' untranslated region (UTR) and a short Viral Priming Protein (VPg)[33]. The virions are approximately 30 nm in diameter and are non-enveloped[33]. The P1 polyprotein, which is encoded by the ORF, is cleaved post-translationally into four structural proteins (VP1-4)[52]. The viral capsid consists of 60 copies of VP1-4. This gives the capsid an icosahedral structure with a canyon in VP1 that acts as receptor binding site to the host cell surface receptors[49]. The canyon is 1.2-3.0 nm and surrounds a fivefold vertex of VP1 proteins from adjacent pentamers[34]. The antigenic diversity of RV is attributed to VP1-3. Whereas the VP4 supports the RNA core to the capsid[49]. Non-structural proteins are involved in the replication and assembly of viral genome[49]. The P2 and P3 proteins are processed to produce the VPg, RNA-dependent RNA polymerase and protease[49].

Approximately 100 different RV strains were yielded using clinical specimens from the 1960s and 1970s[53]. Viral capsid-encoding regions, non-coding regions and a limited number of complete genomes were partially sequenced[53]. The original 99 strains were divided into two species: RV-A consisting of 74 serotypes and RV-B containing 25 serotypes[53]. The full genomes of the 99 serotypes have since been sequenced to further aid the understanding of the evolutionary and molecular biology of RV[53]. A novel species, RV-C, was identified and designated by the International Committee on Taxonomy of Viruses in 2009[54]. A classification system based on genetics was developed to classify RV-C. The strains of RV-C have similar genomic organisation to RV-A and RV-B. However, they have several distinct characteristics which support their classification as a new species. A threshold of 13% nucleotide difference in VP1 or at least a 10% difference in VP4/VP2 region (if the sequence of VP1 was unavailable) to identify at least 50 different types of RV-C[55]. Bochkov *et al* were the first to propagate RV-C in vitro using sinus mucosal tissue.

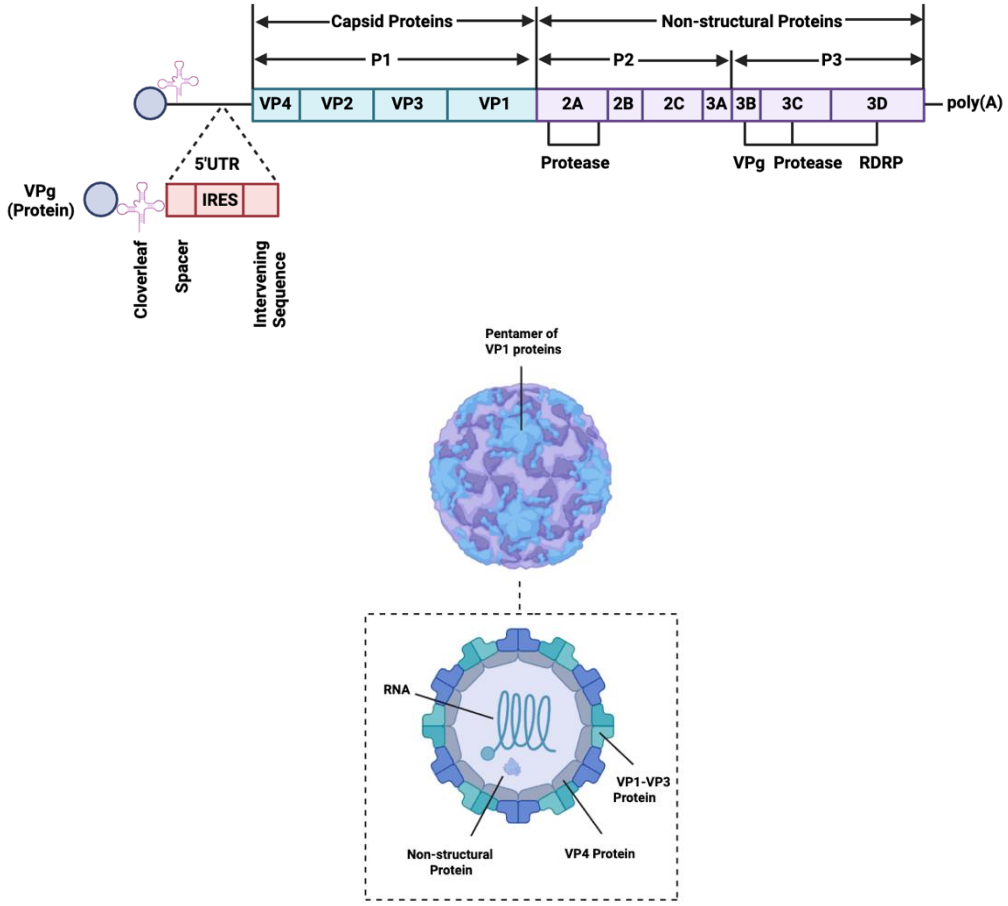


Figure 1-3. Schematic Representation of the Rhinovirus Genome and the Structure of the Rhinovirus Virion. The virion is non-enveloped and composed of VP1-VP4 proteins. Diagram was created using BioRender with Jacobs *et al* as reference.

1.2.1.2 Pathogenesis

Transmission of RV involves via aerosol or person-to-person contact either directly or through a fomite[49]. RVs utilise clathrin-dependent or independent endocytosis or macropinocytosis to enter the cell[49]. The minor group of 12 RV-As bind to the low-density lipoprotein receptor (LDLR) whereas the major group of RV-As and RV-Bs binds to intracellular adhesion molecule-1 (ICAM-1)[51]. Some major group RVs also utilise heparan sulphate as an additional receptor[49]. RV-C binds to cadherin-related family member 3 (CDHR3)[56]. Subsequently, the virions undergo a conformational change which yields

hydrophobic subviral particles[49]. RVs become internalised and are translocated into early endosomes[49]. The mildly acidic pH of the endosomes dissociates RV from the receptor and directs the virus into late endosomes[49]. The viral RNA is released into the cytoplasm through either a pore formed by the viral protein or membrane rupture and is subsequently translated into a polyprotein[49]. Following autocatalytic cleavage into capsid and non-structural proteins, viral polymerase replicates the viral genome[49]. The host cell ribosome translates the RNA into a polyprotein[49]. The progenies are assembled and are released into the nasal cavity where they replicate[51].

RV infection promotes secretion of mucus and vascular leakage which damages the barrier of the airway epithelium[57]. This disruption increases permeability and enables translocation of RV and exposure of receptors such as toll-like receptors (TLR) on the basolateral surface[57]. TLR3 recognises the viral dsRNA whereas TLR7 and TLR8 recognises the viral ssRNA[58]. Upon generation of double-stranded RNA, retinoic acid-inducible gene 1 (RIG-1) and melanoma differentiation-associated gene 5 (MDA5) mediate the response of type I interferon (IFN)[58]. This increases the expression and production of cytokines such as interleukin (IL)-6, IL8, IFN-γ-induced protein 10 (IP-10), epithelial cell-derived neutrophil-activating peptide 78 (ENA78) and regulated upon activation normal T cell expressed and secreted (RANTES)[49]. The production of IFN-β and IFN-γ is activated by interaction with TLR7 and TLR8 which then activates the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway[49]. Secretion of cytokines stimulates recruitment of inflammatory cells such as macrophages, natural killer cells lymphocytes and eosinophils to sites of RV infection[59].

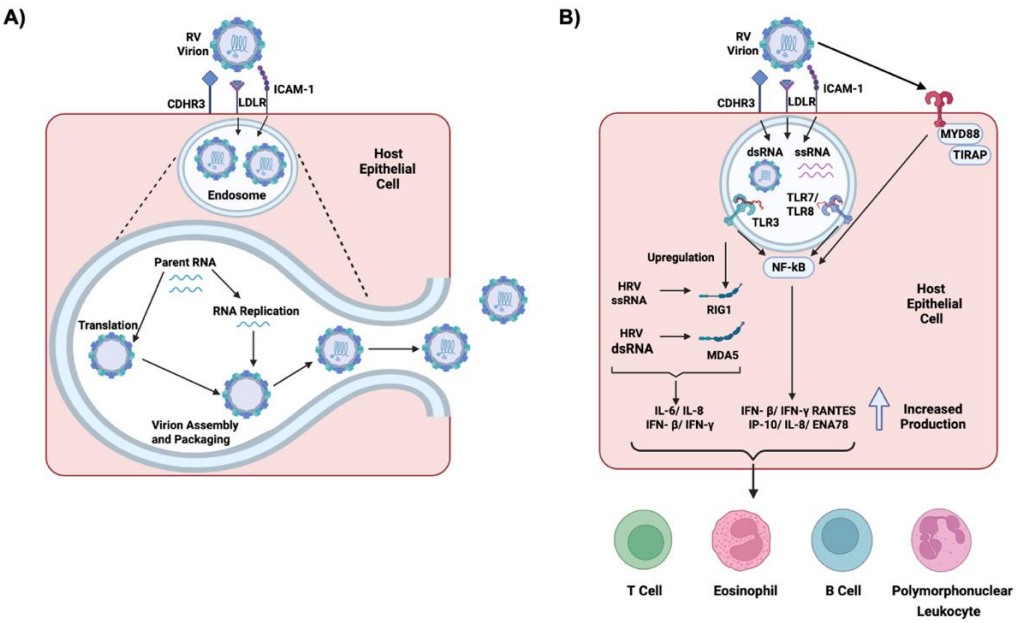


Figure 1-4. RV Infection of the Airway. A) Entry and replication of RV replication. After uncoating, assembly and packaging of the HRV virion commences whilst the parental RNA is translated into capsid and non-structural proteins. **B) Activation of the Immune System.** After recognition of viral ssRNA by TLR7/8 and dsRNA by TLR3, RIG-1 and MDA5 is upregulated. HRV-induced gene expression of IFN stimulates production of cytokine and immune cells. Diagram was created using BioRender with Jacobs *et al* as reference.

1.2.1.3 Impact of RV on Human Health

Common cold, an URTI, is frequently ascribed to RV. Symptoms include sore throat, headaches, sneezing and runny nose[60]. Adults average around 2-3 colds per year with RV responsible for more than half of these cases[33]. Whereas it is estimated that children suffer from colds up to 12 times a year[33]. Although transmission of RV is observed all year around, RV-A and RV-B cases have two peak periods in April/May and September/October[33]. In contrast, RV-C cases tend to peak in winter months. UTRI can cause complications such as acute otitis media (AOM), acute rhinosinusitis (ARS),

and lower respiratory tract infections[60]. AOM, which is the inflammation of the middle ear, can be induced by RV. Following RV infection, an inflammatory process in the nasopharynx is triggered and host immune function is modified[61]. This then enables otopathogens that colonise the nasopharynx to cause harm[61]. Recurrent episodes of AOM can result in acute ear pain, general illness, fever and distress to children and their parents[61]. Though rare, serious complications such as meningitis, acute mastoiditis, and brain abscesses can arise[61]. Globally, 709 million cases of AOM are reported and an estimated 21,000 people die from AOM complications[62].

ARS is an inflammatory disease implicating the nose and the paranasal sinuses[63]. Usually, an inflammatory cascade is initiated by a viral infection[63]. This inflammatory state of the mucosa may enable bacterial infection in few cases[63]. ARS has been found to be attributed to RV in 50% of common cold episodes[64]. ARS poses a substantial problem to public health. Approximately 30 million primary care visits are ascribed to ARS and $11 billion in medical expenditure annually in the United States[65]. Furthermore, ARS is a significant cause of work absenteeism in the United States[63]. Overuse in diagnostic testing and therapies have also been reported[63]. ARS is a common reason for antibiotic prescription, and this raises concern regarding antibiotic resistance[65]. Therefore, further research is required to implement appropriate testing and treatment[63].

Asthma is a respiratory condition characterised by chronic inflammation of the lower respiratory tract[66]. Genetic interactions and environmental stressors contribute to asthma though this is not fully understood[66]. Viral infection, including RV, is thought to exacerbate asthma. Wheezing illnesses attributed to RV infection during infancy or early childhood was found to be the most significant predictor for the development of asthma at age 6[67]. Upon RV infection, proinflammatory mediators are stimulated by the host immune response[68]. This activates neural pathways to augment hyperresponsiveness and obstruction of the airways[68]. T helper cells, a subset of T lymphocytes that express cluster of differentiation 4 (CD4), can be divided into T helper type 1 (Th1) and Th2 and they produce Th1-type cytokines and Th2-type cytokines[69]. Th1-type cytokines mainly produce proinflammatory responses to kill

intracellular parasites or disseminate autoimmune responses[69]. In contrast, Th2-type cytokines such as IL4, 5 and 13 are associated with the stimulation of immunoglobulin E (IgE) and eosinophilic responses in atopy and IL10 with anti-inflammatory response[69]. In healthy individuals, the immune response to RV infection is usually of the Th1 phenotype[68]. In contrast, RV-induced asthma exacerbation is further characterised by an enhanced Th2 immune response and neutrophilia of the airways[68]. Moreover, allergy has been regarded as Th2 weighted imbalance[69]. Asthma is associated with substantial economic burden, mortality, and morbidity[70]. In the US, 7.5% of adults are affected by asthma[71]. Asthma has resulted in 1.8 million hospitalisations and 10.5 million visits to the physician each year[71].

Currently there are no licensed therapeutics for RV infection. Treatments are only supportive. Anti-viral drugs that have been proposed have shown a lack of potency in humans. Although RV infection has a significant impact on public health, no vaccine for RV currently exists. The diversity of RV serotypes poses a significant challenge in developing a vaccine that provides a broad immunity[72]. Protection from RV infection is associated with elevated levels of serotype-specific antibodies[68]. IgA and IgG have been observed in the airways and serum respectively[73]. However, little cross-reactivity is stimulated by neutralising antibodies due to the antigenic heterogeneity[68]. Furthermore, the mechanisms underlying induction of these humoral responses are poorly understood. Approaches such as combining immunogens with Th1-promoting adjuvants provide opportunities for advancing vaccine development[68]. However, a deeper understanding of adaptive response to RV infections is needed and further characterisation of RV-C is required to expedite vaccine development[72].

1.2.2 Respiratory Syncytial Virus

RSV belongs to the *Paramyxoviridae* family and the genus *Pneumovirus*[34]. The virus was first isolated from a chimpanzee with coryza in chimpanzee[34]. The agent was recovered from children with respiratory illness[34]. RSVs are a major cause of lower respiratory tract infections in children and are associated with significant mortality and morbidity[74]. Clinical manifestations can vary from upper respiratory tract infections to pneumonia, bronchiolitis, viral-induced wheeze, and asthma exacerbation[75]. An incidence of 33 million cases and 199,000 deaths in children below the age of 5 to RSV infection have been reported globally[76]. It is estimated that 90% of infants will experience RSV infection in the first two years of their life based on a study conducted in United States[77]. Although rarer in adults, RSV infections are known to afflict the elderly with 5-10% of RSV infection cases in elderly population occurring in nursing homes in the United States[78]. RSV infection in childhood has been associated with the development of asthma at later age. Increases in asthma prevalence in early adulthood have been reported in individuals with severe early RSV bronchiolitis who have experienced recurrent wheezing[79].

1.2.2.1 Structure and Genotypes/Serotypes

RSVs are a negative-sense single-stranded enveloped RNA virus from the *Paramyxoviridae* family[74]. The structure is composed of a helical nucleocapsid which contains fifteen 220 nt negative-sense ssRNAs that are tightly bound to a nucleoprotein[34]. The virus particles are surrounded by a lipid bilayer which contains the G glycoprotein, F fusion protein and SH small hydrophobic protein[34]. G protein mediates attachment of the virus to the host whereas fusion protein is involved in viral fusion and host cell entry[74]. In addition, the virions contain the nucleocapsid (N) protein, phosphoprotein (P) and 22 K matrix proteins M2-1[74]. The negative-strand RNA is encapsidated by the N protein[34]. The large polymerase (L) protein, P phosphoprotein, and cysteine-rich V protein is associated with nucleocapsid[34]. Furthermore, three non-structural proteins NS1, NS2 and M2-2 are produced by RSV-infected cells. The two RSV groups were distinguished originally on the basis of antigenic differences in the attachment of G protein[34]. Through genetic and phylogenetic analysis of the gene sequences of G protein this separation was later

confirmed. Furthermore, this enabled further divisions of the serotypes into genotypes[34]. Group A has 11 genotypes whereas group B has 23 genotypes[33]. Both groups and the different genotypes co-circulates however one serotype is dominant within a given season. It has been hypothesised typically that this serological and genetic diversity contributes to the lack of prolonged immunity and re-infection throughout life.

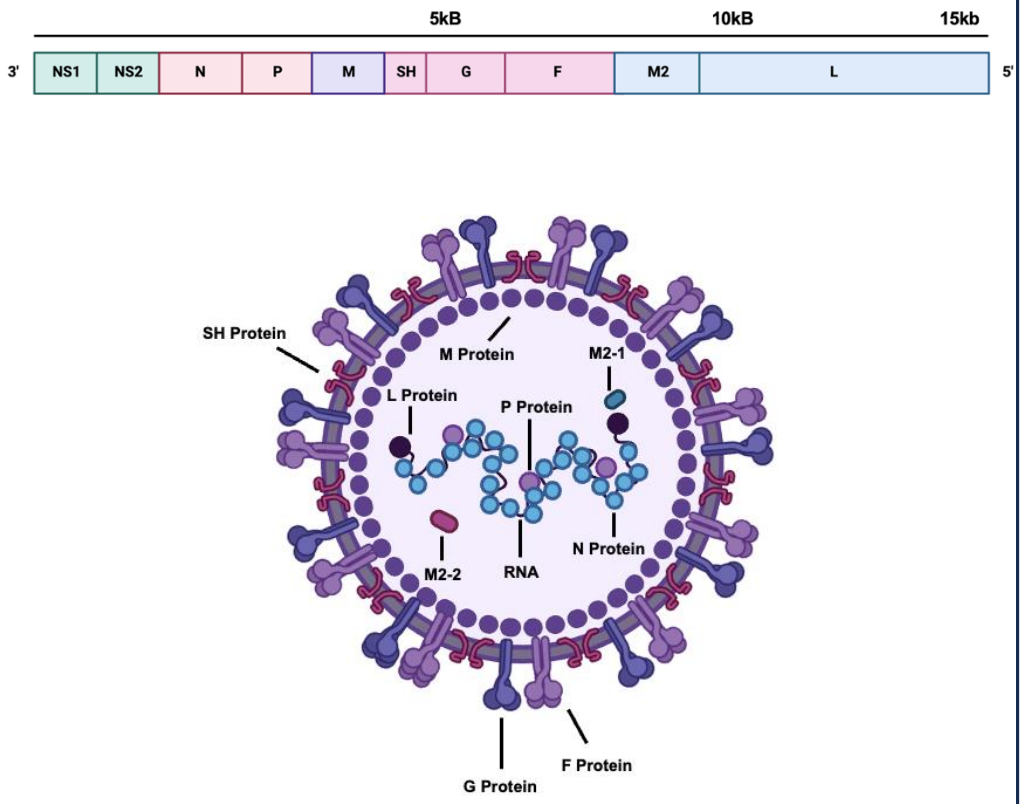


Figure 1-5. Schematic Representation of the RSV Genome and the Structure of the RSV Virion. The virion consists of G glycoprotein, F fusion protein, SH protein, N nucleocapsid protein, L large polymerase protein, P phosphoprotein, M2-1 and M2-2 protein. Diagram was created using BioRender with Boncristiani *et al* as reference.

1.2.2.2 Pathogenesis

RSV is transmitted through direct contact with contaminated air droplets, coughs, and sneezes[80]. Using the G protein, the virus binds to the CX3C motif chemokine receptor 1 CX3CR1[34]. Nucleolin has also been suggested as an alternative receptor for RSV[81]. It was proposed that RSV binds to heparan sulphate proteoglycans (HSPG) to enable interaction between F protein and nucleolin[81]. The virus and the cell membrane fuse together via the binding of F protein to cell GTPase RhoA[34]. The internalisation process of RSV is pH-independent and can happen either in the plasma or the endosomal membrane[34]. Protein F is expressed on the cell surface of infected cells. This promotes fusion between adjacent cells and results in syncytia formation[34]. This is a characteristic of a paramyxovirus cytopathic effect[34]. The RNA is transcribed by the coordinated action of the L, N, P and M2-1 proteins[33]. The resulting set of mRNAs direct the synthesis of the viral proteins with transcription occurring in a gradient with decreasing amounts of transcripts being synthesised from the 3' to the 5' end of the genomic template[33]. NS1 and NS2 are transcribed at the highest abundance from the 5' end. They are involved in the suppression of the innate immune response of infected cells[33]. The polymerase complex then switches from mRNA transcription to viral-RNA replication to generate RNA of polarity opposite to the genome called the anti-genome[34]. The M2-2 protein is partly regulated by this switch[33]. The anti-genome acts as template for the polymerase complex to synthesise the negative-sense RNA genome of the progenies[34]. The plasma membrane buds off to release the mature virus particles. The M-1 and M2 proteins are involved in regulation of transcription and replication whereas the phosphoprotein is essential for functional polymerase[82].

RSV replicates in respiratory epithelia and infects ciliated columnar epithelial cells of the lower and upper airway epithelium[76]. The virus spreads cell-to-cell throughout the respiratory tree and reaches the bronchioles in 1-3 days after the onset of symptoms. Following infection of the epithelial cells by RSV, induction of the inflammatory response results in ciliary dysfunction and apoptosis[83]. Accumulation of debris and the narrowing and oedema of the airways diminishes lung compliance and gives rise to symptoms[83]. The RSV-

infected cells assume a rounded morphology, and they slough onto the smaller bronchioles of the lower airway[82]. Oedema of the surrounding submucosa occurs due to increased mucus secretion and the intrusion of rounded cells[82]. Air becomes trapped in the peripheral areas, and the air absorption causes atelectasis[82]. The mucosal plugs, sloughing of epithelial cell debris, accumulation of macrophages and lymphocytes, necrosis of ciliated cells, formation of syncytia, peribronchiolar and impairment in secretion clearance causes the airways to obstruct[34]. This obstruction and lung hyperinflation is typical of bronchiolitis[33]. In response to the infection, release of proinflammatory cytokines, chemokines and interleukins causes inflammation of the airway and tissue damage[82]. This secretion of cytokines may also induce sensitisation to allergens and promote atopic asthma[74]. RSV-induced pneumonia with eosinophilic cytoplasmic inclusion of epithelial cells, giant multi-nucleated cells and interstitial mononuclear infiltrate frequently co-exists with bronchiolitis[34]. Young babies have immature airways which are unable to compensate for RSV-induced damage[34]. This makes them particularly susceptible to severe RSV disease[34].

Although naturally acquired immunity to RSV is short-lived and incomplete, the severity of disease is reduced upon re-infection. The development of RSV disease is influenced by the pre-existing RSV-specific maternal antibodies. Furthermore, secretary IgA has shown to be correlated with protection. Cell-mediated immune response is key for infection recovery with immunosuppressed individuals being at risk for severe RSV disease and fatal outcome. A major factor in the development of wheezing and asthma exacerbation is the type of immune response mounted upon RSV infection. A deviation towards Th1 response results in effective viral clearance and a milder diseases. Whereas Th2 cytokine response has shown to be associated with more severe illness. A pre-existing deficiency for Th1 response may serve as a marker for pre-disposition to wheezing and asthma in later life.

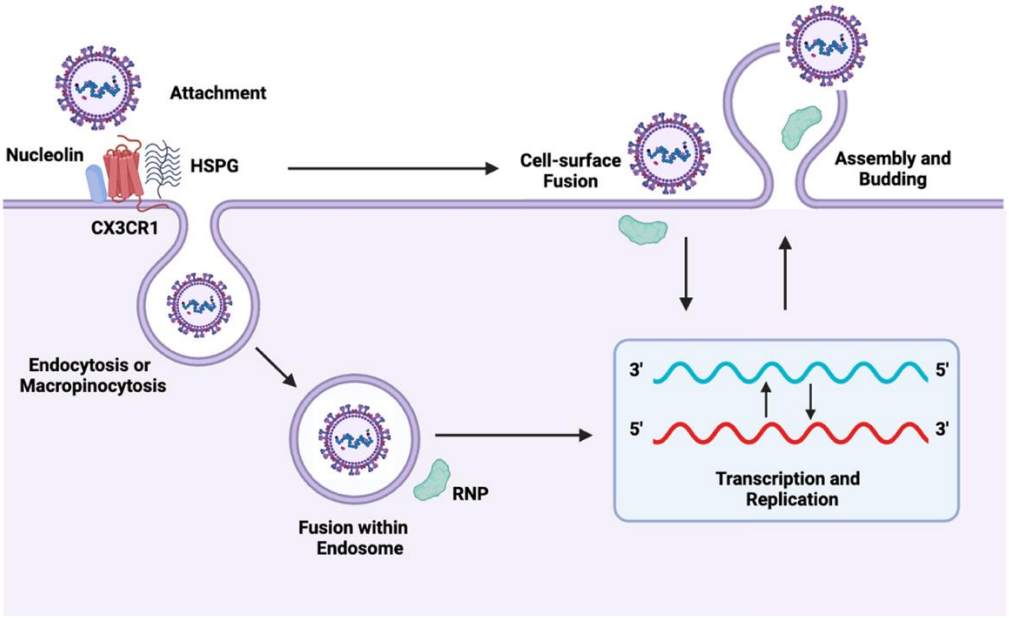


Figure 1-6. RSV Infection of the Airway. RSV binds to CX3CR1 and is internalises via either endocytosis or macropinocytosis. Fusion occurs within the endosome and syncytia is formed. RNA is transcribed and viral proteins are translated. The plasma buds and mature viral particles are released. Diagram was created using BioRender with Boncristiani *et al* as reference.

1.2.2.3 Impact of RSV on Human Health

RSV-induced bronchiolitis is one of the leading causes of hospitalisations for infants[84]. The infection is self-limited but may progress into respiratory failure in infants[83]. By the age of 2, 11-15% of infants suffer from bronchiolitis[83]. At least 5 hospitalisations for every 1000 children under the age of 2 occur depending on disease severity in the United States[83]. In Italy, the cost for acute hospitalisation of bronchiolitis is approximately 2000 euros per patient with an increase to around 8000 euros if paediatric intensive care unit (ICU) is required[84]. Furthermore, RSV aetiology of bronchiolitis was found to contribute to higher cost of hospitalisation due to longer hospital stay and more frequent admission to ICU[84]. Therefore, development of cost-effective interventions against RSV are imperative. Pneumonia in children is frequently caused by RSV[85]. An estimated 3.2 millions episodes of severe RSV pneumonia that

required hospitalisation and 94,600-149,400 deaths was reported globally between 2017-2020[86]. Furthermore, pneumonia is a leading cause of mortality in children with RSV being the most frequent single cause RSV infection has also been associated with extensive mortality and morbidity in critically ill adult patients with severe pneumonia[87]. Age-dependent association between childhood asthma and RSV infection has been found[88]. Children who were not infected with RSV in the first year of life had a substantially reduced risk of developing asthma[88]. Significantly higher rates of asthma and disease severity during childhood have been associated with RSV hospitalisation[89]. RSV infection has been suggested to elicit airway hyperactivity by augmenting a Th2-type response and dampening IFN-γ antiviral immunity[90].

Mid-December to early February is the peak of spread in temperate climates. Treatments for RSV have largely been supportive[33]. Changes in climatic conditions and the resulting influence on behaviour such as increased indoor crowing can also influence the seasonality of RSV infection[33]. Furthermore, a 6-month waning period of protective antibodies has been suggested to impact susceptibility in the following season[33]. Appropriate delivery of supplemental oxygen and hydration is imperative in cases requiring treatments dampening[33]. However, the treatment for bronchiolitis is more extensive. Supportive treatment for infants with RSV bronchiolitis includes bronchodilators such as alpha- and beta-adrenergics, nebulised epinephrine and anti-cholinergics, and prevention of electrolyte imbalance and hypoxemia[34]. However, there is no conclusive evidence that bronchodilators have a positive effect on the disease outcome. Currently the only anti-viral drug approved for infants with RSV is ribavirin[33]. The drug is delivered by small-particle aerosol or a ventilator. However, ribavirin clinical trials have lacked power to provide definitive recommendation[34]. Nevertheless the anti-viral is recommended for infants and young children with underlying health conditions such as cystic fibrosis, congenital heart disease and immunosuppression[34]. Ribavirin use in adults is limited to severely immunocompromised individuals[33].

Passive immunisation with humanised monoclonal antibodies is used to protect high-risk infants during seasonal outbreaks. One such mAb, palivizumab, is administered every month during five months in the first RSV season to limit severe LRTI in children[91]. The mAb targets the F protein and has been market-approved for more than 20 years[92]. However, the high costs and limited efficacy means that the drug is limited to high-risk infants such as immunocompromised infants or pulmonary and cardiac disease. Furthermore the monthly intramuscular injection is arduous and can still allow significant breakthrough infections to occur at low trough antibody levels[92]. A second generation mAb called motavizumab is in phase II clinical trials. The mAb was found to significantly reduce RSV-associated inpatient and outpatient burden in a clinical trial study consisting of 2569 RSV-infected infants[93]. Motavizumab forms a complex with a 24-residue peptide. This structure corresponds to the epitope on the F protein[94]. Modelling studies have suggested that motavizumab recognises a different quaternary configuration of F protein than the one observed in a homologous structure[95]. This gives motavizumab greater potency than palivizumab[95].

Various factors have hindered the development of a RSV vaccine. Deficiencies in the innate and adaptive immune response of neonates makes infantile vaccination difficult[96]. Furthermore, maternal antibodies may subdue the development of intrinsic antibody responses[96]. Significant advances in vaccine development have been made, and two vaccines have been approved: GSK's Arexvy™ (RSVPreF3 OA) and Pfizer's Abrysvo™ (RSVPreF3). Arexvy™ contains recombinant RSV F protein which is stabilised in the prefusion confirmation. This exposes epitopes that are targeted by neutralising antibodies. Previously mice and macaques were immunised with stabilised prefusion forms of the F protein[97]. This elicited significantly greater neutralising activity than the postfusion form[9797]. A phase 1-2 study was conducted to investigate several formulations of the RSVPreF3 OA vaccine including with and without an AS01-based adjuvant[98]. The vaccine was associated with an acceptable safety profile and induced neutralising antibody responses in older adults[98]. Elevated RSV-specific CD4+ T-cell frequencies were observed in older adults to similar levels to those observed in young adults after

vaccination[98]. $AS01_E$ formulations were less reactogenic than $A01_B$ formulations which contain twice the dose of immunostimulants[98]. Therefore, $AS01_E$-adjuvanted formulation was chosen for further development[98]. A randomised, placebo-controlled, ongoing trial (AReSVi-006) was conducted in 17 countries in Africa, Asia, Australia, Europe and North America[99]. Individuals aged 60 years or over were followed for at least two consecutive seasons in the Southern Hemisphere and three consecutive RSV seasons in the Northern Hemisphere[99]. The efficacy of the vaccine was 82.6%. Furthermore, the efficacy was 94.1% against severe RSV-associated lower respiratory tract disease (LRTD)[99].

Pfizer's Abrysvo™ is a bivalent, unadjuvanted vaccine which provides broad protection against RSV lower respiratory tract disease[100]. The vaccine is composed of stabilised prefusion F protein from RSV A and B antigenic subgroup[100]. Variations of the F protein sequence of RSV-A and B subgroups are clustered in antigenic sites making it a major target for potent neutralising antibodies[100]. The vaccine was advanced to RSV vaccine efficacy study in older adults immunised against RSV disease (RENOIR) phase III trial[100]. The randomised, double-blind, placebo-controlled trial was conducted in adults ≥60 years of age from 240 sites across Argentina, Canada, Finland, Japan, Netherlands, South Africa and the United States[100]. The participants received a single intramuscular infection of the vaccine at a dose of 120 µg or a placebo[100]. The vaccine was shown to prevent RSV-associated LRTD and RSV-associated acute respiratory disease without any evident safety concerns[100]. The efficacy was 66.7% against RSV-associated LRTI with at least two signs or symptoms, 85.7% against LRTI with at least three signs or symptoms, and 62.1% against RSV-associated acute respiratory illness[100]. Nevertheless, the need to develop an RSV vaccine for infants still remains. Furthermore, identification of biological targets for therapies and continued development of prophylaxis and anti-viral therapeutics is important[90]. The impact of the protection conferred by RSV vaccine in early childhood on the development of asthma later in life is also essential to understand[90].

1.2.3 SARS-CoV-2

Severe acute respiratory syndrome-corona virus 2 (SARS-CoV-2) is the causative agent of COVID-19 which is responsible for the global ongoing pandemic[101]. According to the World Health Organisation (COVID-19) Dashboard, more than 7 million deaths have been confirmed as of 7th September 2025. SARS-CoV-2 belongs to the *Betacoronavirus* genus. The genus also includes SARS-CoV responsible for the 2022-2004 SARS epidemic), Middle East respiratory syndrome coronavirus (MERS-CoV), human coronavirus (HCoV)-OC43 and HCoV-HKU1[102]. The virus is transmitted through respiratory droplets expelled during face-to-face exposure to coughing, sneezing, and talking[101]. The severity of SARS-CoV-2 infection can range from mild, severe to critically ill[103]. The mortality rate for patients admitted to ICUs falls between 40-50% and is elevated in patients who have developed acute respiratory distress syndrome (ARDS)[103]. Moreover, elderly, immunocompromised or individuals with chronic comorbidities are at higher risk of developing severe COVID-19 disease[103]. The COVID-19 pandemic has had a considerable socio-economic effect which has impacted a wide variety of industries such as healthcare, education, agriculture, and finance[104]. The pandemic has resulted in a global economic crisis with a loss of approximately 90 trillion USD in the global economy[105].

1.2.3.1 Structure

SARS-CoV-2 is a positive-sense single-stranded RNA virus. The 26000-30000 bases long genome is encased inside a capsid[106] and contains two large overlapping ORFs (ORF1a and ORF1b). Four structural proteins (S spike protein, E envelope protein, M membrane protein, N nucleocapsid protein), and nine accessory proteins are encoded by the genome. ORF1a and ORF1b are processed further to generate sixteen non-structural proteins (NSP1-NSP16). Assembled as a homotrimer, S protein is inserted into the membrane in multiple copies giving it a crown-like appearance. Upon infection, the S protein is cleaved into S1 and S2 subunits. The S1 subunit folds into four domains during the pre-fusion confirmation: N-terminal domain (NTD), receptor binding domain (RBD), and two C-terminal domains (CTD1 and CTD2). The pre-fusion S2 subunit is encompassed by these four domains to

form a central helical bundle with heptad repeat 1 (HRP1) structural motif bending towards the viral membrane. Three RBDs make up the apex of the S protein with two distinct conformations: ‘down’ for receptor-inaccessible state and ‘up’ for receptor accessible. Conformational change is initiated during the post-fusion state to disengage the S1 subunit form S2. The S2 subunit goes through a cascade of refolding events to form an elongated stable trimer.

A central protein of SARS-CoV-2, the E protein forms a homopentameric cation channel that is essential for pathogenicity[107]. The M protein consists of two transmembrane domain-swapped three-helix bundles and two intravirion domains to form a mushroom-shaped dimer[108]. The protein also possesses a hinge region which is highly conserved and integral for conformational changes. Using the positively charged intravirion domain, M protein facilitates the recruitment of components for viral assembly. In addition, M protein promotes membrane curvature, membrane budding and binding of nucleocapsid[109]. An integral component of virions, the N protein binds to viral genomic RNA to mediate packing of the RNA into a ribonucleoprotein complex[110]. Additionally, other roles include transcription and replication of viral RNA, immune regulation and cytoskeleton organisation[110]. The N protein has two NTDs and CTDs which are connected by a central linker region (LKR)[110]. The NTD has shown to be responsible for RNA binding whereas CTD is involved in both dimerization and RNA binding[111].

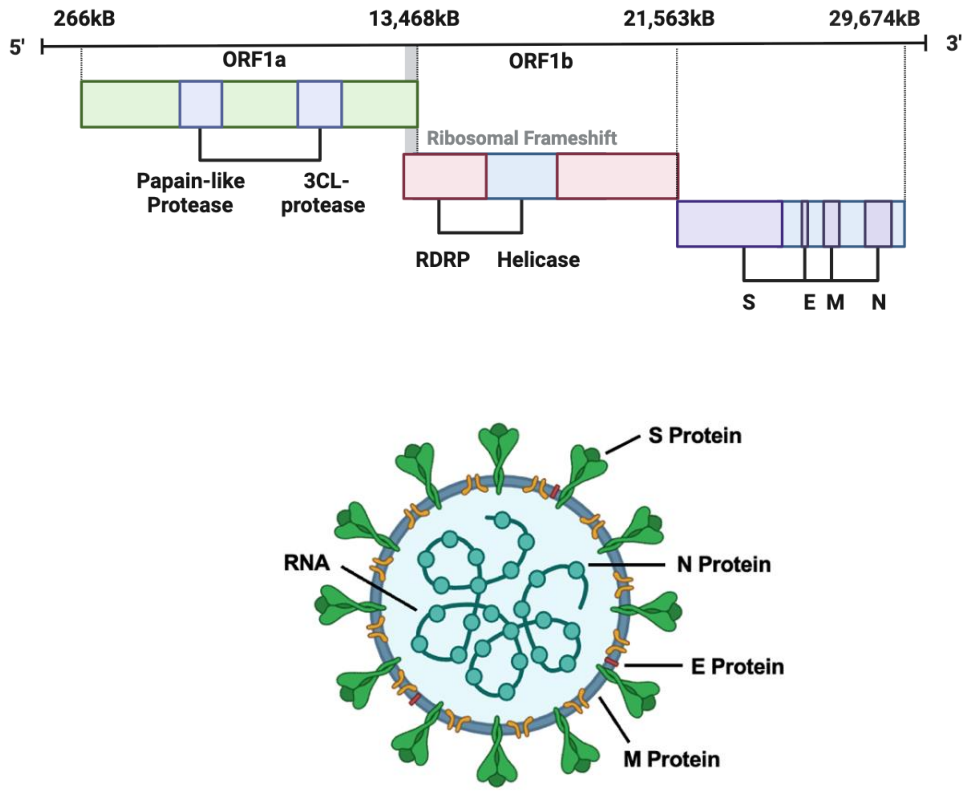


Figure 1-7. Schematic Representation of the SARS-CoV-2 Genome and the Structure of the SARS-CoV-2 Virion. The genome contains the ORF1a and ORF1 which are processed to generate non-structural proteins. The virion is composed of E envelope protein, which is essential for pathogenicity, M membrane protein which facilitates viral assembly, N nucleocapsid protein which mediates packing of RNA, and S spike protein which is involved in viral entry and pathogenicity. Diagram was created using BioRender with Boncristiani *et al* as reference.

1.2.3.2 Pathogenesis

SARS-CoV-2 primarily invades ciliated and goblet epithelial cells and binds to angiotensin-converting enzyme 2 (ACE2). ACE2 is a carboxypeptidase which is essential to the renin-angiotensin-aldosterone system. Additionally, several molecules have been suggested to serve as alternative receptors such as L-SIGN (CD209-L) and tyrosine-protein kinase receptor UFO (AXL). Amraei *et al* reported that CD209-L interacts with the spike receptor-binding domain with

N92 N-glycosylation of the CD209-L being occupied[112.] Whereas study by Wang *et al* found that AXL interacts specifically with the N-terminal domain of S protein to promote SARS-CoV-2 entry[113]. Upon ACE2 binding, S1 subunit undergoes a conformational change to a metastable state to overcome the natural repulsion between the host cell membrane and the virus. This transition is mediated by two proteolytic cleavage: one at S1-S2 boundary by furin and other at S2' site. The cleavage of the S1 subunit subsequently exposes the S2' cleavage site of the S2 subunit. S2' cleavage is dependent on the entry pathway taken: one via the endosomes and other through the host cell surface. This is dependent on the target cell proteases in particular transmembrane serine protease 2 (TMPRSS2) and cathepsin L. If either insufficient TMPRSS2 is expressed by the host cell or the SARS-CoV-2-ACE2 complex does not encounter TMPRSS2, the complex is internalised into endolysosomes via clathrin-mediated endocytosis. Here the cathepsins perform the cleavage of S2' in the presence of an acidic environment. In contrast, the cell surface is the site of S2' cleavage if TMPRSS2 is present.

During both pathways, the fusion peptide (FP) is exposed following S2' site cleavage. S1 dissociates from S2 which induces a conformational change in the S2 subunit particularly in HRP1. This propels the FP forward into the target membrane to initiate membrane fusion. A fusion pore is formed following fusion between the viral and cellular membrane at the C-terminus to release viral RNA into the host cell cytoplasm for replication. Following the release of viral genomes, the virus hijacks the host cell machinery to facilitates transcription and translation of viral proteins. Viral replication accelerates in the later stages of the infection which can compromise the integrity of the epithelial-endothelial barrier[114]. Infection of epithelial and the pulmonary capillary endothelial cells enhances inflammatory response and influx of cytokines and chemokines[114].

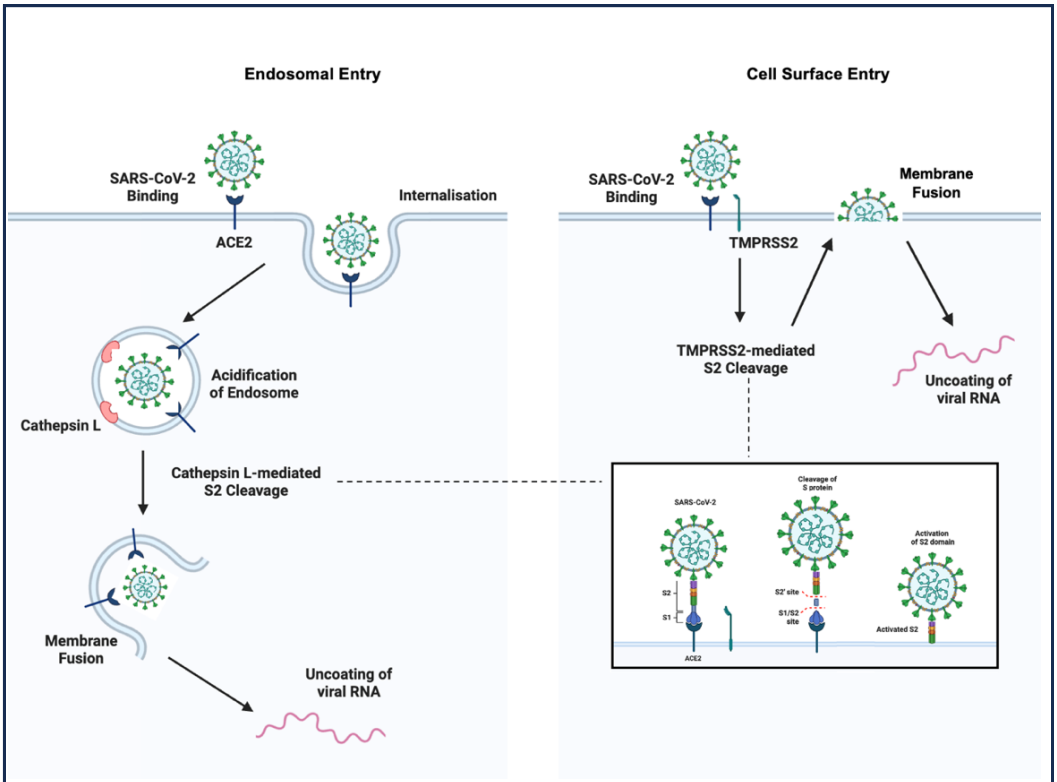


Figure 1-8. SARS-CoV-2 Infection of the Airway. Diagram was created using BioRender with Jackson *et al* as reference. In endosomal entry pathway, SARS-CoV-2 binds to ACE2 and is internalised through clathrin-mediated endocytosis. Cathepsins perform the S2 cleavage. Whereas in cell surface entry pathway, SARS-CoV-2 binds to ACE2, and S2 cleavage occurs in the presence of TMPRSS2. Following S2 cleavage, fusion peptide is exposed and conformational change is induced. Fusion between the viral and membrane results in the formation of a fusion pore. RNA is released into the cytoplasm for uncoating and replication. Viral genomes are released. The host cell machinery is hijacked by the virus to stimulate the transcription and translation of viral proteins. Diagram was created using BioRender with Jacobs *et al* as reference.

1.2.3.3 Impact of SARS-CoV-2 on Human Health

COVID-19 represents a wide spectrum of clinical manifestations from asymptomatic to critical illness[115]. Critical illness is characterised by acute respiratory failure with the need for mechanical ventilation, septic shock, and failure of multiple organ systems[115]. COVID-19 represents a wide spectrum of clinical manifestations from asymptomatic to critical illness[115]. Common symptoms of SARS-CoV-2 include cough, fever, fatigue, loss of smell and/or taste, and dyspnoea[116]. Gastrointestinal symptoms such as nausea, vomiting, diarrhoea, and abdominal pain can occur in some patients[117]. The expression of ACE2 across a variety of different brain regions such as brain stem, cerebral cortex, choroid plexus, striatum, middle temporal gyrus, paraventricular nuclei of the thalamus, and posterior cingulate gyrus could explain the different neurological manifestations presented by COVID-19 patients[118]. These include headaches, dizziness, impairment of the smell, stroke, intracerebral haemorrhage[118]. Furthermore, individuals with severe epilepsy were found to have an increased likelihood of developing severe complications[119]. COVID-19 can result in numerous complications including lung fibrosis, cardiac thrombosis, and inflammation, arterial thromboses, and mood dysfunctions[120]. Regarding the respiratory system, complications such as dyspnoea, dependence on ventilation and oxygen, abnormalities of pulmonary function tests can occur[120].

Critical illness is characterised by acute respiratory failure with the need for mechanical ventilation, septic shock, and failure of multiple organ systems[115]. Acute respiratory distress syndrome (ARDS) is a dangerous complication of SARS-CoV-2 infection. ARDS is defined by new or worsening respiratory failure coupled with hypoxemia and infiltrates of the alveoli[121]. Upon SARS-CoV-2 infection, the epithelial and endothelial structures are damaged resulting in increased permeability and oedema[122]. This can lead to diffuse alveolar damage consisting of capillary congestion, formation of hyaline membranes, hyperplasia of type 2 pneumocyte, necrosis of pneumocytes and alveolar oedema[123]. ARDS can also trigger proliferation of fibroblast and myofibroblast which can lead to pneumonia[122]. As one of the most frequent complications of severe COVID-19 disease, progression into ARDS is one of

the highest predictors of COVID-19 mortality[121]. The mortality rate of individuals with COVID-19-induced-ARDS ranges from 23-56%[105]. Treatment requires longer hospital stays, advanced medical equipment and therapies[105]. Risk factors include advanced age, comorbidity with heart disease, diabetes mellitus, COPD, and coagulopathy disorder, smoking, obesity, immunosuppressive medication and ethnicity[105].

Long COVID is a multisystemic illness that can be developed following COVID-19 infection. During long COVID symptoms continue or develop 3 months after SARS-CoV-2 infection and persist for at least 2 months[124]. In most cases the symptoms will resolve however neuropsychiatric symptoms can last for more than 6-12 months[124]. These include brain fog, sleep disorders, cognitive disorders, depression, and anxiety[125]. Type 2 diabetes, cardiovascular, cerebrovascular, and thrombotic disease, chronic fatigue syndrome and dysautonomia are some of the complications arising from long COVID[126]. Alteration of T cells such as exhausted T cells, reduction in the number of CD4+ and CD8+ effector memory cell numbers, and increased expression of Programmed Cell Death Protein 1 (PD1) on central memory cells for at least 13 months were reported in individuals with long COVID[126]. Persistence of SARS-CoV-2 reservoir could potentially drive symptoms of long COVID[126]. Viral proteins and RNA were detected in lung tissues, cardiovascular system, brain, lymph nodes, reproductive system, eyes, hepatic tissue, breast tissue and plasma in patients with long COVID-19[125].

The pandemic has led to an unprecedented acceleration of the vaccine development process and treatment strategies[127]. Traditional vaccine development may take 15 or more years due to a long discovery phase, pre-clinical experiments, toxicology studies, clinical trials, and large-scale production[128]. The knowledge gained from initial vaccine development for SARS-CoV and MERS-CoV helped accelerate the process[128]. Following interim analysis of phase I/II clinical trials data, rapid progression to phase III was initiated[128]. RNA-based, DNA-based, non-replicant viral vector, and protein subunit COVID-19 vaccines have been developed[129]. BNT162b2 (Comirnaty®; BioNTech, Pfizer)[130] and mRNA-1273 (Moderna)[131] are lipid nanoparticle-encapsulated mRNA vaccines. The ChAdOx1nCoV-19 vaccine

(AZD1222; AstraZeneca) comprises a replication-deficient chimpanzee adenoviral vector ChAdOX1 which houses the spike glycoprotein[132]. Whereas NVX-CoV2373 (Novavax) is a recombinant nanoparticle vaccine that contains that prototype strain of the spike glycoprotein and Matrix-M adjuvant[133]. Furthermore, small-molecule antiviral drugs such as remdesivir, nirmatrelvir-ritonavir and molnupiravir and eleven monoclonal antibodies have been marketed to treat COVID-19[127]. Immunomodulatory drugs such as the glucocorticoid dexamethasone, Janus kinase inhibitor barticinib and cytokine antagonist tocilizumab may help treat hospitalised patients with critical or severe COVID-19[127]. Effective strategies for treating COVID-19 are still required including treatments that tackle drug resistance conferred by emerging variants[127]. It is also important to develop pan-coronavirus inhibitors and combination therapies to target variants and prepare for future outbreaks of coronaviruses[127]. Furthermore, the mechanisms underlying long COVID still need to be elucidated and there is no drug currently on the market that has been approved for treating long COVID[127].

1.3 Host Genetic Factors Associated with Respiratory Viruses

Host genetics plays an important role in susceptibility and severity of viral infections as indicated by twin studies[134]. Single nucleotide polymorphisms (SNPs) of host genetic factors have been associated with viral infections[135]. The variations range from substitution of amino acids to formation of splice variants[135]. Some of the SNPs have consequences on the function such as production of non-functional proteins and could potentially influence disease susceptibility and severity[135]. Several SNPs associated with RV, RSV and SARS-CoV-2 infection have been identified including variants of host receptors. Studies range from genomic-wide association studies (GWAS) to case-studies of individuals with a specific polymorphism. It is important to note the differences in cohort sizes and population groups between studies when generalising results. Although GWAS have primarily focused on population of European descent, global population are increasingly being represented[136]. To prevent false positive that may be caused by differences in allele frequencies between population groups, GWAS have implemented different methods to control for ethnicity[137]. Principal component analysis (PCA), the standard tool for correcting population stratification, analyses genetic data onto axes to visualise population structure[136]. The principal components (PCs) is computed within the cohort or estimated from an external reference such as 1000 Genomes Project[136]. The PCs can be subsequently included as covariates in linear regression model[137]. Another method for controlling population stratification is the mixed-model approach which utilises the genetic relationship matrix (GRM) to model the genetic relatedness between every participant in the study[137]. A greater statistical power may be yielded using the mixed-model due to increased sample size and control of variance explained by genetic relatedness[136].

1.3.1 Rhinovirus

1.3.1.1 IL6

Upon infection and tissue injury, the pro-inflammatory cytokine interleukin- 6 (IL6) is produced by immune cells[138]. IL6 stimulates immune reactions, acute phase response and haematopoiesis to support the host immune defences[138]. RV infection upregulates IL6, and it is thought to contribute to inflammation of the airways[139]. The IL6 gene is located at chromosome 7 (7p21-24)[138]. Polymorphisms of cytokine genes can result in the variation in the level of cytokine produced[140]. The -174G/C of rs1800795 SNP is a G-to-C two-point substitution at position 174 in the 5' flanking region of the IL6 gene[141]. The rs1800795 SNP alters the expression and functional activity of IL6[141]. Doyle *et al* explored the influence of rs1800795 SNP on rhinovirus illness[142]. Following genotyping of seventy-two adults and exposure to rhinovirus, infection was tracked[142]. The low IL6 production phenotype of rs1800795 was found to be significantly associated with higher severity of rhinovirus symptoms[142]. The SNP was significantly associated with throat symptoms ($p = 0.037$) and general symptoms ($p = 0.04$)[142]. Therefore, the polymorphism may predict severity to rhinovirus illness.

1.3.1.2 IL10

Interleukin-10 (IL10) is an anti-inflammatory cytokine produced by different cell types such as monocytes, T helper cells, dendritic cells, and macrophages depending on the tissue type, time stage of the infection and the stimulus response[143]. IL10 suppresses antigen presentation and limits the production of inflammatory mediators in macrophages and monocytes[143]. Furthermore, IL10 serves an essential role in B and T cell biology[143]. Th2 differentiation has shown to be augmented by RV infection, and this is regulated by IL10 induction[144]. Encoded by the IL10 gene located on chromosome 1 (1q31-31), the cytokine is a 178 amino acids long protein[143]. Few polymorphisms of IL10 exist including the G-to-A substitution polymorphism rs1800896 which is located in a negative putative binding site. A negative putative binding site is a site where transcription factors binds to and acts as repressors to decrease the rate of gene transcription[145]. A case-control study by Holster *et al* from Finland found that the AG or GG genotype of the rs1800896 polymorphism,

which is associated with high production of IL10, was found less often in infants with rhinovirus bronchiolitis compared to controls ($p = 0.009$)[146]. The AT or TT genotype of the rs1800890 polymorphism was also less present in infants with rhinovirus bronchiolitis compared to healthy control ($p = 0.03$)[146]. Helminen *et al* also found that infants with rhinovirus bronchiolitis were less likely to carry the G allele of rs1800896 polymorphism in a study consisting of 139 infants with bronchiolitis and 400 control subjects from Finland[147]. The rs1800896 and rs1800890 polymorphism may play a role in RV infection susceptibility.

1.3.1.3 NOS2A

Nitric oxide (NOS) has important roles in various physiological processes in multiple organs including the lungs such as mucus transport, neurotransmission, and surfactant production[148]. Expression of NOS2A mRNA and production of NOS has been shown to increase following RV infection[149]. NOS genes are encoded by three genes: neuronal NOS by NOS1 on chromosome 12, inducible NOS by NOS2A on chromosome 17, and endothelial NOS by NOS3 on chromosome 3[150]. Few studies have explored the role of NOS2A polymorphism on rhinovirus infection. Alvarez *et al* observed significant association between rs1060826 NOS2A polymorphism and bronchiolitis attributed to rhinovirus ($p < 0.05$)[151]. The cohort consisted of 181 infants diagnosed with severe bronchiolitis and 536 healthy adult controls from Brazil[151].

1.3.2 RSV

1.3.2.1 ADAM33

A disintegrin and metalloproteinase 33 (ADAM33) is a member of a group of zinc-dependent metalloproteinases encoded by the ADAM33 gene located in chromosome 20 (20p13)[152]. ADAM33 is thought to be involved in airway remodelling and stimulation of cell proliferation and differentiation[153]. Polymorphism of ADAM33 has shown to be associated with RSV disease severity. A study by Drysdale *et al* found that the C allele of the rs2787094 polymorphism of ADAM33 is associated with an increased risk of developing RSV lower respiratory tract infection (LRTI) ($p = 0.04$)[154]. The sample size consisted of 148 infants with RSV infection and 119 controls from UK[154]. The polymorphism may increase the risk of developing severe RSV disease.

1.3.2.2 CCL5

Chemokine (C-C motif) ligand 5 (CCL5), also known as RANTES, is a chemokine that plays an important role in immune response against viral infections by binding to G protein-coupled receptors (GPCRs) such as CC chemokine receptor type 5 (CCR5)[155]. As a chemokine, CCL5 stimulates recruitment and migration of leukocyte, T cells, NK cells, mast cells, dendritic cells, and basophils[156]. RSV infection has been reported to induce production of CCL5 and recruit RSV-specific cells which in turn promotes additional CCL5 production[157]. Polymorphism of the CCL5 gene has shown to be associated with RSV susceptibility and severity. Two polymorphisms in the CCL5 gene, -403G/A of rs2107538 and -28C/G of rs2280788, are found in the promoter region[158]. The In1.1T/C genotype of rs2280789 polymorphism is found in the intronic regulatory element that facilitates transcription of CCL5[158]. The frequency of -403G/G genotype of rs2107538 ($p = 0.017$, OR = 0.49), -28G/G genotype of rs2280788 ($p = 0.010$, OR = 0.33), and In1.1T/C genotype of rs2280789 ($p = 0.037$, OR = 0.52) polymorphisms of the RANTES gene was reported by Hattori *et al* to be lower in Japanese infants with severe bronchiolitis than controls[159]. The study population consisted of 59 infants diagnosed with severe RSV bronchiolitis and 201 control subjects[159]. Moreover, the alleles -403A ($p = 0.012$, OR = 0.56), -28G ($p = 0.032$, OR = 0.45) and In1.1C ($p = 0.015$, OR = 0.56) were lower in infants with severe RSV

bronchiolitis[159]. The -403A and -28G alleles are associated with higher promoter activity[159]. Therefore lower frequency of the alleles suggest that infants with lower production of CCL5 might be less susceptible to severe RSV bronchiolitis. These results suggest that risk of severe RSV infection may be conferred by CCL5 polymorphisms.

1.3.2.3 CCR5

CCR5 is a GPCR involved in the regulation of trafficking and effector functions of immature dendritic cells, memory and effector T-lymphocytes and macrophages[160]. CCR5 is also a co-receptor for the cellular entry of R5 strains of human immunodeficiency virus 1 (HIV-1) and 2 (HIV-2)[160]. Located on the short arm of chromosome 3, the receptor is encoded by the CCR5 gene[161]. The first instance of CCR5 polymorphism with regards to RSV infection was documented by Hull *et al* in 2003[162]. Family-based and case-control genetic-association analysis was conducted on 580 infants with RSV-induced bronchiolitis from UK and both of their parents[162]. Case-control analysis revealed that the -2554T allele of the rs2734648 SNP was found more frequently in the RSV group than controls ($p = 0.01$, OR = 1.25)[162]. In addition, the -2459G allele of rs1799987 SNP was more commonly found in the RSV group than the control group ($p = 0.02$, OR = 1.22) with a stronger effect found in individuals homozygous for these alleles[162]. Polymorphisms of CCR5 may make individuals more susceptible to RSV infection.

1.3.2.4 CD14

Cluster of differentiation 14 (CD14) is a membrane receptor protein for toll-like receptors and mediates innate immunity responses to tissue injury and pathogens in monocytes and macrophages[163]. The protein has been implicated in various diseases such as infectious diseases such as sepsis, metabolic diseases, autoimmune conditions like amyotrophic lateral sclerosis, cancer, and cardiovascular disease[163]. CD14 has also been shown to mediate innate immune response to RSV[164]. The gene encoding for CD14 is located on chromosome 5[165]. A number of polymorphisms in the CD14 gene have been identified including the rs5744455 variant. A study by Inoue *et al* observed that the C allele ($p = 0.034$) and the C/C genotype ($p = 0.037$) of the -550 C/T rs5744455 polymorphism were significantly increased in children with RSV-

induced bronchiolitis compared to control subjects[166]. 54 children with RSV bronchiolitis and 203 control individuals from Japan made up the sample size[166]. Furthermore, the serum level of soluble CD14 was significantly elevated in children with CC genotype of rs5744455 polymorphism compared to CT and TT genotypes. This suggests that CD14 polymorphism may be involved in the development of RSV bronchiolitis.

1.3.2.5 CNR2

Cannabinoids and endocannabinoids activate the cannabinoid receptors 1 (CB1) and 2 (CB2)[167]. Neuroimmune interactions are modulated by the CB1 expressed in the brain and CB2 expressed in the periphery[168]. CB2 is encoded by the CNR2 gene which is located on chromosome 1 (1p36.11)[169]. The gene is composed of six exons which spans a 90 kb region. Tahamtan *et al* investigated the effect of CNR2 polymorphism on RSV infection. The rs35761398 polymorphism was analysed in 90 inpatient and 90 outpatient children with acute respiratory tract infection (ARTI)[170]. Children carrying the Q allele had a two-fold increase in the risk of developing severe ARTI following infection with RSV ($p = 0.026$, $OR = 2.148$)[170]. CNR2 polymorphism may confer risk to severe RSV infection.

1.3.2.6 CXC3R1

CX3CR1 is a 7-transmembrane G-protein coupled receptor that binds to its natural ligand fraktaline and the CX3C motif of the RSV G-protein[171]. CXC3R1 are expressed in immune cells, and they recognise and enter inflamed tissue to initiate the innate immune system[172]. The rs3732378 (T280M) polymorphism of CX3CR1 results in a substitution from threonine to methionine on position 280 resulting in reduced fraktaline binding and antiviral immune responses. A study of 82 RSV-infected children and 126 healthy control subjects from Greece by Amanatidou *et al* reported a significant association between rs3732378 SNP and risk for severe RSV bronchiolitis[173]. RSV-infected children exhibited predominance of the T/M or M/M genotype compared to the control ($p = 0.025$, $OR = 2.03$)[173]. Compared to controls, the T/M or M/M genotype were significantly higher in RSV-infected individuals with oxygen saturation of less than 92% ($p = 0.006$, $OR = 3.8$) and individuals requiring oxygen ($P = 0.019$, $OR = 2.2$)[173]. The authors hypothesised that the

rs3732378 polymorphism may increase affinity for RSV G protein and enable viral entry by inhibiting the fraktaline and antiviral immune response further.

1.3.2.7 HLA

The major histocompatibility complex (MHC) is a large complex with an essential function in the immune system[174]. The MHC encodes for glycoproteins called the human leukocyte antigens (HLA) which are essential for immune response induction and regulation, and the selection of the T cell repertoire[174]. Positioned on the short arm of chromosome 6, the MHC gene family is made up of three groups: MHC class I, II and III[175]. Class I and II are highly polymorphic and are involved in antigen processing and presentation[176]. The classical HLA-A, B, C and the non-classical MHC-Ib genes E, F and G make up the class I genes[176]. Whereas class II is comprised of genes that encode α and β chains of the class II molecules including DPA1, DPB1, DQA1, DQB1, DRA, and DRB1 genes[176]. Few studies have explored the influence of HLA polymorphism on RSV infection. Salas *et al* conducted a whole exome sequencing of 54 children with RSV diagnosis and control group from the 1000 Genomes Project[177]. They reported an association between rs1047985 HLA-DQA1 ($p = 2.21 \times 10^{-10}$) and RSV susceptibility[177]. Moreover, rs935 HLA-DPB1 ($p = 3.51 \times 10^{-8}$) was also associated with susceptibility to RSV infection[177]. These results suggest that HLA polymorphism may induce susceptibility to RSV disease.

1.3.2.8 IFN-γ

Interferon gamma (IFN-γ) orchestrates several protective functions in response to infections[178]. IFN-γ exerts its immunomodulatory effect by augmenting antigen processing and presentation, promoting an anti-viral state, and heightening leukocyte trafficking[178]. The IFN genes are mapped onto chromosome 9 (9q21). Some studies have investigated the effect of IFN-γ polymorphisms and RSV infection. Gentile *et al* conducted a case-control study consisting of 77 infants with RSV infection and 107 control subjects from the United States[179]. The rs2430561 of IFN-γ was associated with severity of LRTI ($p < 0.001$) and duration of ICU stay ($p = 0.021$)[179]. The IFN-γ CA microsatellite (rs3138557) was investigated in 218 patients with RSV bronchiolitis and 303 healthy control individuals from China[180]. There was a

lower frequency of CA12 repeats in patients with RSV bronchiolitis compared to controls (p = 0.018)[180]. IFNG polymorphism may influence susceptibility to and severity of RSV infection.

1.3.2.9 IL4

Interleukin-4 is a cytokine that promotes class switching of B-cells towards IgE production and drives differentiation of naïve T cells towards Th2 cells[181]. IL4 levels are also elevated during RSV infection and induce type 2 cytokine production[182]. IL4 and IL13 shares the IL13 receptors which is comprised of IL4Rα binding chain and IL13α1 chain[183]. Several polymorphisms have been identified including the rs2243250 polymorphism which is located upstream of the transcription initiation site of the IL4 gene on chromosome 5 (5q31)[184]. The polymorphism has been shown to enhance transcription and overexpression of IL4[184]. The association between IL4 polymorphisms and RSV infection have been investigated in several studies. In a study population of 200 RSV-infected children and 447 control subjects from the Netherlands, Hoebee *et al* found that the -589T allele of rs2243250, which is associated with increased promoter activity, was significantly higher in RSV group than controls (p = 0.04, OR = 1.43)[185].

Zhang *et al* reported significant differences in frequencies of -590C/T of rs2243250, and -33C/T of rs2070874 in Chinese Han children. The study sample consisted of 218 patients with RSV bronchiolitis and 303 control subjects[186]. The frequency of the T allele of rs2243250 was higher in the RSV group compared to the control group (p = 0.003, OR = 1.69)[186]. Moreover, a gene-region association study of the 5q31 cytokine cluster was conducted by Forton *et al* in 782 RSV cases and 1045 controls of white European ancestry[187]. The following IL4 polymorphisms showed association with increased risk of severe RSV bronchiolitis at the genotype level: rs2243250 (p = 0.05), rs2070874 (p = 0.012), rs734244 (p = 0.042), rs2243270 (p = 0.023)[187]. Whereas association with increased risk of severe RSV bronchiolitis at the allele level was exhibited by the following polymorphisms: rs2070874 (p = 0.049), rs2227284 (p = 0.028), rs2243270 (p = 0.032)[187].

There have been a few studies that did not find a significant association between IL4 polymorphisms and RSV infection. A case-control study by Marr *et al* on 240 Canadian children with severe RSV-induced bronchiolitis and 285 control subjects could not replicate the association of IL4 SNP with RSV disease severity[188]. Furthermore, Puthothu *et al* also could not replicate the association between the -590C/T of rs224325 and severe RSV infection in their cohort of 131 RSV-infected children and 270 control subjects[189]. The differences in population groups could reflect the contradictory nature of the findings in research. Polymorphisms of IL4 receptor have also been explored to see whether they influence susceptibility to or severity of RSV infection. Tapia *et al* used a candidate-gene approach to investigate whether IL4R polymorphisms influence RSV infection in 60 infants with moderate RSV infection and 58 with severe RSV infection[190]. Higher frequency of G allele of rs1801275 was observed in infants with severe RSV infection compared to moderate infection ($p = 0.047$, OR = 1.85)[190]. The rs1801275 SNP influences the structure of the IL4Rα binding chain and signalling pathways[183]. The Ile/Ile genotype of rs1805010 was less frequently found in infants with RSV infection who went on to develop recurring wheezing compared to those who did not develop wheezing ($p = 0.046$, OR = 0.14)[190].

1.3.2.10 IL6

As outlined above on section 1.5.1.1 IL6 stimulates the innate and adaptive immune response. Elevated levels of IL6 have been associated with RSV infection risk[191]. Interestingly, study by Gentile *et al* revealed that low IL6 production rs1800795 polymorphism was significantly associated with severity of RSV infection including the length of hospital stay ($p = 0.009$) and length of oxygen supplementation ($p = 0.025$)[179]. This could be attributed to IL6 inducing expression of nasal and upper respiratory disease but protecting against expression of lower airway disease[179].

1.3.2.11 IL8

The pro-inflammatory cytokine interleukin-8 is produced by a variety of different cells for example macrophages, lymphocytes, fibroblasts, epithelial and endothelial cells. Secretion of IL8 stimulates trafficking of neutrophiles to sites of inflammation. Concentration of IL8 elevates during RSV infection and is linked to disease severity[192]. The IL8 gene comprises of four exons and three introns[193]. Several nuclear factors found on the 5'-flanking region of the IL8 gene synergistically activate the IL8 gene with NF-κB and activator protein 1 (AP-1)/ CCAAT-enhancer-binding proteins (C/EBP)[193]. The rs4073 polymorphism and its association with RSV infection have been reported in several studies. Following genotyping of 117 infants with bronchiolitis and 180 healthy infants, Hull *et al* identified IL-251A allele of the rs4073 SNP at 251bp upstream from the IL8 transcription start site[194]. A significant increase of the frequency of A allele, which is associated with elevated IL8 production, was observed in infants with RSV bronchiolitis from the UK ($p = 0.014$)[194]. Furthermore, 53 out of the 83 infants with rs4073 SNP required oxygen therapy for more than two days and seven required mechanical ventilation[194]. These results suggest that the A allele of rs4073 SNP near the IL8 gene may confer genetic pre-disposition to RSV disease severity[194].

Goetghebuer *et al* reported higher frequency of the rs4073 A allele among children who experienced wheezing post-bronchiolitis compared to the control group ($p = 0.02$)[195]. The study cohort consisted of 134 children with RSV-induced bronchiolitis and 706 control children from the UK[195]. In contrast, a case-control study of 101 RSV-induced pneumonia patients and 101 pneumonia patients by Lu *et al* found that the T allele of rs4073 was higher in the RSV-pneumonia group compared to controls ($p = 0.0002$, OR = 2.08)[196]. The difference in the risk allele could potentially be due to variation in different population groups. Furthermore, the rs4073 polymorphism does not have a direct effect on IL8 function. However, the association with a range of different human diseases suggests that the polymorphism is a marker for functional variability[195]. Fine mapping studies have suggested that the effect on the disease association is located on a region consisting of IL8 and RSS56 gene. RSS56 does not seem to have a known immuno-modulatory function[195].

1.3.2.12 IL10

Interleukin-10 is an anti-inflammatory cytokine which is produced by different cell types such as monocytes, T helper cells, dendritic cells, and macrophages depending on the tissue type, time stage of the infection and the stimulus response[143]. Upon receptor binding, JAK/STAT signalling is activated which results in inhibition of pro-inflammatory cytokines secretion, reduction in antigen presentation and phagocytosis[197]. IL10 is encoded by the IL10 gene on chromosome 1 (1q31-31)[143]. Furthermore, production of IL10 has shown to be induced upon RSV infection[198]. IL10 gene has multiple polymorphism including those that regulate IL10 promoter activity such as rs1800896 (-1082G/A) and rs1800872 (-592C/A). The rs1800896 is located in a negative putative transcription factor binding site and is a G to A substitution[145]. Whereas the rs1800872 SNP is found in a region with a putative negative transcription factor binding site and confers a C to A substitution[145]. Studies have focused on IL10 polymorphism and their effect on RSV infection. A study by Gentile *et al* reported that the low IL10 production genotypes of rs1800896 polymorphism was associated with increased frequency of pneumonia development following RSV infection ($p = 0.02$)[179]. Moreover, Drysdale *et al* reported significant association between the C allele of rs1800872 SNP and parental reported wheeze following RSV-induced LTRIs ($p = 0.045$)[154]. Parental reported wheezing was significantly lower in infants with AA genotype compared to AC genotype ($p = 0.045$)[154]. These results suggest that IL10 polymorphisms may increase the likelihood of developing a severe RSV infection.

1.3.2.13 IL13

The cytokine interleukin-13 (IL13) is an immunoregulatory cytokine which is predominantly secreted by activated Th2 cells[199]. As an important mediator of allergic inflammation pathogenesis, IL13 functions through the IL13Rα1/IL4Rα complex and activates STAT6 to regulate this inflammation[200]. IL13 is implicated in mucus secretion and hyperresponsiveness of the airway following RSV infection[201]. Few studies have inspected the influence of IL13 polymorphisms on susceptibility and severity to RSV infection. The Gln allele of rs20541 was reported by Ermers *et al* to be found more frequently among

children with late wheezing following RSV infection (p = 0.007, OR = 3.27)[202]. The study population consisted of 73 children with RSV-induced LRTI, and 447 controls subjects from the Netherlands[202]. The rs20541 polymorphism encodes an amino acid change from arginine to glutamine on position 130. This amino acid switch may potentially change the affinity towards the IL13Rα1 chain of the IL4/IL13 receptor[202].

A study by Puthothu *et al* reported that -1112C/T genotype of rs1800925 SNP was significantly associated with severe RSV infection (p = 0.026)[189]. The cohort comprised 131 RSV-infected children and control group of 270 individuals from Germany[189]. The -1112C/T genotype of rs1800925 SNP was overrepresented in children with severe RSV disease[189]. The polymorphism is thought to modify IL13 expression of binding of nuclear factors to the IL13 promoter[189]. However, it is important to note that the close proximity of IL4 and IL13 and the similarity in their function means that it is hard to decipher the influence of the polymorphisms from both IL13 and IL4 on RSV disease[189]. Furthermore, a study by Choi *et al* found no association between IL13 -1512A/C of rs181457, -1112C/T of rs1800925 and G2044A of rs20541 and RSV infection in the study population of 105 RSV-infected children and 315 control subjects[203].

1.3.2.14 IL15

Interleukin-15 (IL15) is essential for the development, differentiation, and survival of natural killer cells[204]. IL15 modulates innate and adaptive immunity by facilitating the development of protective and inflammatory immune response to pathogens[204]. Via viral replication and NF-κB dependent mechanisms, production of IL15 is increased following RSV infection[205]. IL15 is encoded by the IL15 gene found on the 4q31 chromosome[206]. The gene spans 35 kb and is composed of nine exons and eight introns[207]. A study by Janssen *et al* found that the T allele of rs2254514 showed association with RSV susceptibility (p = 0.0260)[208]. The cohort consisted of 470 children with RSV-induced bronchiolitis and 1008 controls from the Netherlands[208].

1.3.2.15 IL17

Interleukin-17 (IL17) is a proinflammatory cytokine that is involved in immune response and inflammation[209]. IL17 stimulates the release of other cytokines, chemokines, and antimicrobial peptides from myeloid and mesenchymal cells[209]. A significant increase has been reported in IL17 production during RSV infection[210]. Furthermore, this increase has been positively associated with severity of pneumonia (Long). IL17 gene is positioned on chromosome 6[211]. Polymorphism of IL17 gene has also been associated with RSV susceptibility in literature. Association between G/G genotype of the rs7747909 polymorphism and RSV susceptibility was reported by Janssen *et al* ($p = 0.0455$)[208].

1.3.2.16 IL18

Pro-inflammatory cytokine interleukin-18 (IL18) stimulates the immune response and enhances Th1 response by inducing gamma interferon production[212]. Bronchiolitis induced by RSV has also been shown to be accompanied by elevated nasal IL18 response[213]. The IL18 gene consists of seven exons and is located on chromosome 11. Several polymorphisms of IL18 gene exist including the intronic variant rs360721. A study by Puthothu *et al* examined the influence of various IL18 polymorphism on RSV infection[214]. The genotype analysis of 154 children with severe RSV infection and 270 control subjects from Germany was conducted[214]. A significant association between 133C/G of the rs360721 SNP and severe RSV infection was found ($p = 0.043$)[214].

1.3.2.17 JARID2

Jumonji and AT-rich interaction domain containing 2 (JARID2) are histone demethylases that are critical for recruiting epigenetic regulators called the polycomb repressive complexes (PRCs)[215]. PRCs are responsible for haematopoietic stem and progenitor cell function[215]. The JARID2 gene is located on the short arm of chromosome 6[216]. In comparison to case-control studies focusing on cytokine genes, the influence of JARID2 on RSV infection has been relatively understudied. However, a whole exome sequencing study

by Salas *et al* of 54 identified an association between rs72837837 of JARID2 and RSV susceptibility ($p = 7.59 \times 10^{-07}$)[177].

1.3.2.18 JUN

The oncogene Jun encodes the protein c-Jun which is part of the activation protein 1 (AP-1) family of transcription factors[217]. c-Jun is involved in various processes such as apoptosis, progression of the cell cycle and differentiation[217]. Upon expression, c-Jun forms dimers with another AP-1 protein called c-Fos to form AP-1[218]. Furthermore, c-Jun has also been shown to protect human tumor cells from DNA damage-induced apoptosis[217]. Only a few studies have looked at the impact of JUN polymorphisms on RSV infection susceptibility and severity. Janssen *et al* reported a strong association between A allele of rs11688 ($p = 0.0093$) and G/A genotype ($p = 0.0285$) and RSV bronchiolitis has been reported[208].

1.3.2.19 MARCO

Macrophage receptor with collagenous structure (MARCO) is part of the class A scavenger receptor family[219]. Expressed on the surface of macrophages, MARCO binds to unopsonised pathogens to enable ingestion by macrophages[219]. The MARCO gene has been mapped to chromosome 2 (2q14)[220]. A study by High *et al* looked at the influence of MARCO polymorphism in a case-control study consisting of 189 infants with severe RSV bronchiolitis and 109 infants with mild RSV bronchiolitis from Argentina. A significant association between rs1318645 polymorphism and a greater risk of RSV disease was reported ($p = 0.017$)[221].

1.3.2.20 MMPs

Matrix metalloproteinases (MMPs) constitute a family of proteases which disintegrate the extracellular matrix during angiogenesis, tissue remodelling and metastasis[222]. They cleave major components of the extracellular matrix such as collagen, casein, gelatin, and elastin[223]. A study by Drysdale *et al* investigated polymorphism of MMPs on RSV susceptibility and severity. It was reported that the rs2664349 and rs2664352 polymorphisms of MMP16 were associated with reduced lung function test following RSV infection. The gene locus for MMP16 is mapped onto chromosome 8 (8q21.3)[224]. MMP16, has a

role in epithelial-to-mesenchymal-transition (EMT) and migration of neural crest cells[222]. Moreover, MMP16 has been shown to affect pulmonary expression and may play a role in alveoli development[225].

1.3.2.21 MUC4

Mucin-4 (MUC4) is membrane-bound mucin that is deregulated in cancer[226]. The mucin is expressed in numerous cancers such as lung, colon, stomach, and pancreatic cancer[227]. MUC4 is encoded by the MUC4 gene found on the chromosome 3q29[227]. Whole exome sequencing study by Salas *et al* reported association between RSV susceptibility and polymorphisms of the MUC4 gene: rs201623571 (3.55×10^{-10}), rs529417345 (9.40×10^{-10}), rs548345415 (9.40×10^{-10})[177].

1.3.2.22 MX1

Myxoma resistance 1 (MX1) is an interferon-induced GTP-binding protein which has been shown to mediate influenza A virus infection[228]. MX1 interacts with influenza ribonucleoprotein complex and perturbs the interaction between polymerase protein 2 and nucleoprotein to restrict with the assembly of the complex[229]. Only a few studies have explored the effect of MX1 polymorphism on RSV infection susceptibility and severity. In a case-control study of 374 infants with RSV infection and 305 healthy controls from Argentina, individuals with the A allele of rs469390, which is associated with increased RSV G protein expression, had a higher risk of developing severe disease compared to those with the G allele ($p = 0.021$, OR = 1.750)[230].

1.3.2.23 NCF2

Neutrophil cytosolic factor 2 (NCF2) gene consists of 16 exons with 40 kb and is located on chromosome 1 (1q25)[231]. NCF2 is implicated in different cancers such as oesophageal squamous cell carcinoma[231] and hepatocellular carcinoma[232]. The NCF2 gene is located on chromosome 1 (1q25)[231]. Janssen *et al* reported an association between NCF2 polymorphism and RSV infection susceptibility[208]. The AA genotype of rs2274064 polymorphism was associated with susceptibility to RSV bronchiolitis ($p = 0.0154$)[208] .

1.3.2.24 NOS2A

As mentioned in section 1.5.1.3, NOS2A is involved in various processes of the airways. A study by Drysdale *et al* observed that TT genotype of rs1060826 NOS2A was associated with reduced lung function and T allele was associated with increased days of coughing symptom[154]. Janssen *et al* reported strong association between A allele of rs1060826 ($p = 0.0031$) and G/A genotype ($p = 0.0117$) RSV susceptibility[208].

1.3.2.25 OPRM1

μ-opioid receptor (MOR) is one of the most common opioid receptors. Following a painful stimulus, endogenous opioids are released and bind to MOR[233]. Therefore, they mitigate pain sensation in the absence of opioid medication[233]. MOR is encoded by the OPRM1 gene[233]. Found on the exon 1 of OPRM1 gene, the A118G rs1799971 variant is responsible for the amino acid substitution from asparagine to aspartate at position 40 of MOR[234]. Studies on rs1799971 polymorphism have mainly focused on pain and opioid use[233]. Salimi *et al* conducted a case-control study investigating the effect of rs1799971 on RSV infection. There was an association between rs1799971 and RSV disease severity ($p = 0.015$, OR = 0.731)[235].

1.3.2.26 OR13C5

Olfaction plays an important role in numerous aspects of life[236]. The main olfactory epithelium located in the upper region of the nose detects the odorants[236]. The odor molecules interact with olfactory sensory neurons through the odorant receptors[236]. Upon binding of the molecules to the receptors, an electrical signal is transmitted to the olfactory bulb in the brain which results in perception to the odour and emotion and behaviour response[236]. Olfactory receptors are expressed in numerous tissues including those of the lungs and play a role in cell proliferation, migration, and secretion[237]. Polymorphism of olfactory receptors have been relatively understudied with regards to RSV infection. Whole exome sequencing of 54 patients with RSV bronchiolitis by Salas *et al* showed association ($p = 2.53 \times 10^{-12}$) between rs199665292 polymorphism of olfactory receptor 13C5 (OR13C5) gene and RSV susceptibility[177]. Several reports have shown

respiratory viruses can infect the epithelium of the olfactory nerve that connects to the nasal cavity[177]. Although RSV does not pose a threat as a neurovirulent virus, Riel *et al* suggested that the neural route could potentially be used by RSV as a shortcut[238].

1.3.2.27 SNAI3-AS1

The long non-coding RNA SNAI3 Antisense RNA1 (SNAI3-AS1) induces EMT to promote the growth of hepatocellular carcinoma (HCC) and metastasis[239]. SNAI3-AS1 binds to frameshift protein 1, regulated mothers against decapentaplegic homolog 7 (SMAD7) expression and activates the TGF-B/SMAD pathway which affects HCC tumorigenesis[239]. The role of SNAI3-AS1 polymorphism in RSV infection has been reported by Salas *et al*. The rs3794628 polymorphism showed association with RSV susceptibility ($p = 8.52 \times 10^{-7}$)[177].

1.3.2.28 Surfactant Proteins

Surfactant proteins (SP) are calcium-dependent collagenous lectins[240]. SP-A and SP-D are involved in innate immune response, and they employ a diverse range of mechanisms such as opsonisation and aggregation to neutralise viruses[241]. SP-B and SP-C function in metabolism and recycling of surfactant and facilitates the spreading of surfactant across the lining of the alveoli[240]. The genes involved in encoding of SP-A and SP-D are located on the long arm of the chromosome 10q22-q23[240]. Löfgren *et al* investigated the association between SPA polymorphisms and severe RSV infection using 86 RSV-infected infants and 95 control individuals. The frequency of SP-A2 allele 1A3 of rs1965708 was significantly overrepresented in infants with RSV compared to the control group ($p = 0.006$)[242]. Furthermore, the frequency of infants with severe RSV carrying the allele pool with lysine at amino acid 223 instead of glutamine (Q223K) of rs1965708) was higher compared to the control group ($p = 0.023$)[242]. Position 223 is found in the carbohydrate binding domain which facilitates binding between RSV and SP-A. This substitution of glutamine to lysine on position 223 may modify RSV binding and alter susceptibility to RSV infections[242].

Association between reduced lung function and an increase in the A allele of the SFTPC polymorphism was also associated with the number of parental reported visits to the hospital for RSV infection[154]. Few studies have also investigated the association of SP-D on RSV infection. A study by Ampuero *et al* researchers reported that Thr_{11} (C allele) of SP-D rs721917 was significantly more common in RSV-infected infants than control subjects (p = 0.0046) from Chile[243]. Thr_{11}/Met_{11} (CT genotype) of SP-D rs721917 was more commonly found in infants with severe RSV compared to control group (p = 0.037)[243]. Furthermore, Thr_{160} (A allele) of SP-D rs2243639 was significantly more common in the RSV group compared to infants with mild infection (p = 0.018) and control subjects (p = 0.046)[243]. In contrast, a study by Lahti *et al* reported significant increase of Met_{11} allele (rs1059046) in infants with severe RSV[244]. Higher frequency of Met_{11} allele was found in the RSV group compared to controls in a cohort consisting of 84 infants with severe RSV and 93 controls from Finland (p = 0.033)[244]. The amino acid switch in position 11 caused by the polymorphism affects the function of the protein[244]. Found near the amino acids Cys15 and Cys20, the amino acid at position 11 are implicated in the disulphide bonding of the subunits of SP-D molecule[244]. The disulphide bridges are important for stabilisation of SP-D structure and maintenance of antiviral capacities[244]. The rs1059046 polymorphism may disrupt the disulphide bridges and subsequently the SP-D structure[244].

1.3.2.29 TNF

Tumor necrosis factor (TNF) family consists of TNF alpha (TNFα), TNF beta (TNFβ), Fas ligand (FasL), LIGHT, CD40 ligand (CD40L), TNF-related apoptosis inducing ligand (TRAIL)[245]. Numerous pathological and physiological processes are mediated by TNF such as tumor cell necrosis and apoptosis[245]. TNF has an important role in the defence to viral, bacterial, and parasitic infections[246]. RSV infection has been shown to drive disease pathogenesis through necroptosis of alveolar macrophages[247]. Positioned on chromosome 6, TNFα and TNFβ genes are 3 kb and consists of 4 exons[245]. Few studies have looked at the influence of TNF polymorphism on RSV infection. Janssen *et al* documented the association between the C/T genotype and RSV infection susceptibility (p = 0.0363)[208].

1.3.2.30 VCAM-1

Vascular cell adhesion molecule-1 (VCAM-1) is a cell adhesion molecule that facilitates inflammation-induced vascular adhesion and the trans-endothelial migration of leukocytes [248]. VCAM-1 has also been associated with the progression of multiple immunological disorders such as asthma, rheumatoid arthritis, cancer, and transplant rejection[248]. The levels of VCAM-1 are upregulated following RSV infection[249]. Furthermore, association between C allele of rs1041163 and RSV susceptibility has been reported by Janssen *et al* ($p = 0.0469$)[208].

1.3.2.31 Toll-like Receptors

Toll-like receptors are pattern recognition receptors that support the innate immune system through the recognition of pathogen-associated molecular patterns (PAMPs) from pathogens[250]. One of the TLRs, TLR4, is expressed by immune system cells such as macrophages, monocytes, and dendritic cells[251]. TLR4 recognises PAMPs which triggers proinflammatory response[251]. During RSV infection, the RSV F protein binds to the hydrophobic pocket of the myeloid differentiation factor (MD-2)/TLR4 complex to facilitate activation of RSV[251]. Various studies have investigated the association of TLR4 polymorphisms and RSV infection. In a study population of 99 infants with RSV-induced bronchiolitis and 90 healthy adults, Tal *et al* reported that there was a significant overrepresentation of rs4986790 polymorphism in the RSV group compared to the control group ($p = 0.009$)[252]. A significant overrepresentation of the rs4986790 polymorphism was reported in the RSV group compared to infants with mild bronchiolitis ($p = 0.003$)[252]. The polymorphism confers an aspartate-to-glycine substitution in the TLR4 protein[252]. Overrepresentation of the rs4986791 polymorphism, a threonine-to-isoleucine transition at position 399, was also found in infants with severe RSV bronchiolitis compared to infants with mild bronchiolitis ($p = 0.01$) and control group ($p = 0.009$)[252].

Analysis of 105 RSV-infected children and 7092 controls by Awomoyi *et al* identified an overrepresentation of A/G (Asp/Gly) genotype of rs4986790 in RSV group compared to the control group ($p < 0.0001$)[253]. Heterozygosity for the C/T (Thr/Ile) genotype of rs4986791 was frequently found in the RSV group compared to controls ($p < 0.0001$)[253]. The results suggest that heterozygosity of rs4986790 and rs4986791 is associated with increased susceptibility to RSV infection. However, Paulus *et al* reported no association between the A/A, A/G or G/G of rs4986790 and susceptibility to RSV infection in a study population of 131 RSV-infected children and 277 control subjects from South-West Germany[254]. This suggests that different genetic factors of different ethnic groups may contribute to severity of RSV infection. Similar to TLR4, TLR8 recognises pathogens and has importance in mediating innate immunity. TLR8 can also sense RNA degradation products in human phagocytes[255]. The gene encoding TLR8 is positioned within a 56 kb interval on the X chromosome[256]. The G allele of rs2407992 was associated with susceptibility to RSV infection ($p = 0.0433$)[208].

1.3.2.32 Vitamin D Receptor

Low vitamin D level is postulated to be a risk factor for RSV infection due to its role in immune response against viral infections and the low vitamin D production during winter where cases of RSV infection are higher[257]. Vitamin D Receptor (VDR) facilitates the biological function of vitamin D therefore polymorphism may potentially influence vitamin D signalling and RSV infection[257]. Vitamin D has been shown to reduce the inflammatory response towards RSV infection in the airway epithelium[258]. Therefore adequate levels of vitamin D may contribute to less severe RSV disease[258]. Polymorphism of the vitamin D receptor and its association with RSV infection has been the subject of some studies. Janssen *et al* reported an association between rs10735810 and RSV susceptibility ($p = 0.0017$)[208]. The rs10735810 polymorphism is a start codon polymorphism resulting from a C-to-T transition and is found three codons upstream from the second initiation codon (ATG)[259]. This polymorphism contains a FokI site which results in a protein that is three amino acids longer than the protein encoded by major F allele[259].

The relationship between rs10735810 polymorphism and RSV infection was examined by Kresfelder *et al* in 296 RSV-infected infants and 113 healthy infants. The heterozygous C/T genotype of rs10735810 SNP was present among a higher proportion of the RSV group compared to the control group ($p = 0.008$)[260]. The allele frequencies of the minor T allele of rs10735810 were higher in the RSV group in contrast to the control group ($p = 0.006$) suggesting predisposition to RSV disease[260]. These results suggest that the rs10735810 SNP exhibits an association with the South African population and highlights the importance of the VDR in RSV disease. Moreover, Drysdale *et al* reported association between reduced lung function and the T/T genotype of rs10735810 SNP following RSV-induced LRTI ($p = 0.05$)[154].

1.3.3 SARS-CoV-2

1.3.3.1 ABO

The ABO blood group system has an essential role in transfusion and transplantation[261]. Blood group A has been reported to enhance SARS-CoV-2 infection[262]. Increased risk in the need for mechanical ventilation and prolonged ICU admission has been identified in patients with A or AB blood group and a critical ill COVID-19 infection[263]. Seven exons comprise the ABO gene with a 1062 bp coding for a glycotransferase[261]. The Severe COVID-19 GWAS Group found an association with COVID-19 induced respiratory failure and the rs657152 polymorphism positioned at chromosome 9 (9q34.2) ($p = 4.95 \times 10^{-8}$, OR = 1.32). The study cohort consisted of 835 COVID-19 patients and 1255 control subjects from Italy and 775 COVID-19 patients and 950 control subjects from Spain. The distribution of ABO blood groups was skewed among COVID-19 patients with respiratory failure who had mechanical ventilation[264]. Higher risk was reported among individuals with blood group A (OR = 1.45) whereas protective effect was found in blood group O (OR = 0.65)[264]. Niemi *et al* reported association between rs912805253 ABO and SARS-CoV-2 infection[265]. It is important to note however that allele frequencies of the ABO, for example A, B, AB and O, differs across different populations[266]. Therefore, the association could be false-positive due to inflated test statistic created by population stratification. The studies by Severe COVID-19 GWAS Group and Niemi *et al* controlled for population stratification.

1.3.3.2 ACE

First identified as an important part of the renin-angiotensin system, Angiotensin converting enzyme (ACE) regulates electrolyte balancing, converts angiotensin I to angiotensin II for the purpose of vascular remodelling and degrades the peptide bradykinin to regulate blood pressure[267]. ACE is encoded by the ACE gene which is composed of 26 exons[268]. Following splicing of exon 13, the mature transcript is composed of 25 exons[268]. Located at 17q23.3 locus, ACE gene possesses a functional polymorphism rs4646994 in intron 16 which exists in the form of insertion (I) allele and deletion (D) allele of 289 bp Alu repeat[269]. Hubacek *et al* investigated the effect of rs4646994 on SARS-CoV-2 infection using a cohort of 408 SARS-CoV-2 positive individuals and 2559 control subjects from Czech Republic[270]. A significant increase in the frequency of I/I homozygotes was observed in SARS-CoV-2 patients compared to controls ($p = 0.02$, OR = 1.55)[270]. This difference was shown to be driven by symptomatic subjects which comprised 60% of the case subjects ($p = 0.002$, PR = 1.78). This suggests that individuals with I/I homozygotes are at increased risk of symptomatic COVID-19. Verma *et al* also studied the rs4646994 polymorphism in a study population of 149 patients with mild COVID-19 disease and 120 patients with severe COVID-19 disease from India[271]. The I allele was significantly less frequent in severe COVID-19 patients ($p = 0.004$, OR = 2.59)[271]. Whereas individuals with D/D genotype showed 3.69-fold higher risk of COVID-19 severity ($p = 0.002$).

1.3.3.3 ACE2

ACE2 is the functional host receptor for SARS-CoV-2 and expressed abundantly in a diverse range of cells in different organs[272]. ACE2 plays an important role in the regulation of blood pressure and electrolyte homeostasis[272]. The enzyme catalysis the conversion of angiotensin I to angiotensin II which then facilitates renal sodium reabsorption, vasocontraction, potassium excretion, blood pressure elevation and initiation of pro-fibrotic and inflammatory pathways[272]. The ACE2 gene is located on the X chromosome[273]. As the receptor for SARS-CoV-2, ACE2 and its polymorphism has been the subject for several studies. Mohlendick *et al* genotyped the rs2285666 ACE2 polymorphism in 297 SARS-CoV-2 patients

and 253 SARS-CoV-2 negative patients from Germany[274]. A significant association was found between the G allele (p = 0.02) and the G/G genotype (p = 0.02) and SARS-CoV-2 infection susceptibility[274]. Carriers with the G/G genotype had a three-fold increase in developing a serious SARS-CoV-2 infection (p = 0.0020). Horowitz *et al* reported that the expression of ACE2 was downregulated by 37% ($p = 2.7 \times 10^{-8}$) and the risk of SARS-CoV-2 infection was decreased by 40% ($p = 4.5 \times 10^{-13}$) by the 190509934 polymorphism[275]. These results suggest that the ACE2 polymorphism influences the susceptibility and severity of SARS-CoV-2 infection.

1.3.3.4 ACSL6

Long-chain acyl-coenzyme A synthetase 6 (ACSL6) is a member of an enzyme family that helps convert long-chain fatty acids into their active form[276]. ACSL6 is essential for the maintenance of docosahexaenoic acid-containing glycerophospholipids[276]. The ACSL6 gene is mapped onto chromosome 5 (5q31)[277]. Polymorphism of ACSL6 and its influence on SARS-Cov-2 infection have been explored in some studies. A GWAS study by Kousathanas *et al* revealed a genome-wide association between rs56162149 and critical COVID-19 ($p = 7.65 \times 10^{-11}$, OR = 1.2)[278]. The study cohort consisted of 7491 critically ill individuals with COVID-19 and 48400 controls[278].

1.3.3.5 APOE

Apolipoprotein E (APOE) is a multifunctional protein involved in the assembly and metabolism of lipoproteins[279]. APOE functions in injury repair of the brain and so APOE polymorphisms have been investigated in studies involving neurodegenerative disorders[280]. APOE has been shown to interact with ACE2 and inhibit the cellular entry of SARS-CoV-2 and reduce inflammation in patients with COVID-19[281]. The APOE gene is located at 19q13 and is polymorphic[282]. The gene encodes one of the three common alleles ε2, ε3 and ε2 from a combination of two non-synonymous polymorphisms rs429358 and rs7412[282]. This gives rise to three homozygous genotypes (ε2ε2, ε3ε3, ε4,ε4) and three heterozygous genotypes (ε2 ε3, ε2ε4, ε4ε4)[282]. Espinosa-Salinas *et al* explored the association between SARS-CoV-2 infection and rs7412 polymorphism. The study population consisted of 50 individuals with SARS-CoV-2 infection, 58 individuals with uncleared SARS-CoV-2 infection and 141

with no infection[283]. A significant association between the T allele of rs7412 and SARS-CoV-2 infection was found (p = 0.039, OR = 0.216)[283]. The allele was found to play a protective role against the infection.

1.3.3.6 ATP11A

ATPase phospholipid transporting 11A (ATP11A) is a phospholipid transferase and regulates phospholipid asymmetry to maintain cell signalling, vesicular transportation, and apoptosis[284]. Polymorphism of ATP11A has been associated with COVID-19 severity. Kousathanas *et al* reported a genome-wide association between critical COVID-19 and rs9577175 ($p = 3.71 \times 10^{-11}$, OR = 1.2)[278].

1.3.3.7 BCL11A

B-cell lymphoma/leukaemia 11A (BCL11A) is a zinc-finger protein that acts as a transcriptional repressor[285]. BCL11A is essential in the development of the brain and hematopoietic system and switching of foetal-to-adult haemoglobin[285]. BCL11A is also involved in the differentiation of leukocytes[278]. The BCL11A gene is located on chromosome 2[286]. Kousathanas *et al* reported a genome-wide association between rs1123573A and critical COVID-19 ($p = 9.85 \times 10^{-10}$, OR = 1.1)[278]. This suggests that the BCL11A polymorphism may confer a genetic predisposition to COVID-19 severity.

1.3.3.8 CCHCR1

Coiled-coil alpha-helical rod protein 1 (CCHCR1) is a protein implicated in hyperproliferative skin disorders such as psoriasis and eczema[287]. The expression of CCHCR1 has been shown to be altered on psoriatic skin compared to healthy skin[287]. The CCHCR1 gene is found within the major psoriasis susceptibility locus PSORS1[287]. Association between CCHCR1 and severity towards COVID-19 infection have been reported in literature. Genome-wide association between rs143334143 and critical COVID-19 infection was identified by Pairo-Castineira *et al* ($p = 8.8 \times 10^{-18}$, OR = 1.8)[288].

1.3.3.9 DPP9

Dipeptidyl peptidase 9 (DPP9) possesses peptidase and enzymatic activities that have been associated with several diseases such as cancers[289]. DPP9 has been documented to modulate signalling cascades to regulate cell proliferation and survival[289]. An increase in DPP9 expression has been found in patients with COVID-19[290]. Several studies have looked at the influence of DPP9 on RSV infection. Genome-wide significant association between the risk allele A of rs2109069 on chromosome 19p13.2 and critical COVID-19 has been reported ($p = 4.99 \times 10^{-8}$, OR = 1.4)[288]. The serine protease encoding DPP9 is implicated in inflammasome activation and so may be involved in SARS-CoV-2-induced inflammatory lung injury[288]. Niemi *et al* reported genome-wide association between rs2109069 and increased risk of severe COVID-19 symptoms ($p = 2.76 \times 10^{-17}$)[265]. These results suggest that variants of DPP9 may confer susceptibility to or severity of SARS-CoV-2 infection.

1.3.3.10 EFNA4

Ephrin A4 (EFNA4) acts as an oncogene in several cancers including hepatocellular carcinoma where the protein is highly expressed[291]. Overexpression of EFNA4 has been shown to contribute to the growth of lung tumour cells and poor prognosis of lung cancer[292]. In severe cases of COVID-19 the expression of EFNA4 has shown to be upregulated[293]. EFNA4 is encoded by the E4NA4 gene which is located on chromosome 1[294]. Polymorphism of EFNA4 has been reported to be associated with critical COVID-19 disease. Kousathanas *et al* identified a genome-wide association between rs114301457 and critical COVID-19 ($p = 6.8 \times 10^{-10}$, OR = 2.4)[278].

1.3.3.11 ELF5

E74-like factor 5 (ETS domain transcription factor) is a member of the ETS transcription factors which are involved in cellular proliferation and tumorigenesis[295]. ELF5 has important regulatory roles including determination of trophectoderm cell fate, development of the mammary gland and differentiation of mammary luminal progenitor cells[296]. ELF5 is expressed in multiple organs including the lungs, kidney, stomach, mammary gland, prostate and the bladder[295]. Gene encoding ELF5 is mapped onto

chromosome 11 (11p13-15)[295]. A GWAS study identified a genome-wide association between rs61882275 and critical COVID-19 ($p = 1.61 \times 10^{-10}$, OR = 1.1)[278]. This suggests that the polymorphism may increase the risk of developing severe COVID-19 disease.

1.3.3.12 FOXP4

Forkhead box transcription factors are implicated in several differentiation and developmental processes in particular mediation of neuronal and lung cell differentiation and the development and function of T cells[297]. Forkhead box P4 (FOXP4) has been shown to be induced upon exposure to a thermogenic stimuli and regulates the thermogenic functionality of adipocytes[297]. The FOXP4 gene is localised on chromosome 6. Few studies have looked at the influence of FOXP4 polymorphisms on SARS-CoV-2 infection. A genome-wide association between increased risk of severe COVID-19 symptoms and rs1886814 FOXP4 has been reported ($p = 1.11 \times 10^{-9}$)[265].

1.3.3.13 FUT2

Galactosidase 2-alpha-L-fucosyltransferase 2 (FUT2) is implicated in secretory status of ABO antigens. The FUT2 gene is positioned on the 19q13.33 chromosome and comprises a 6865 bp intron between 2 exons of 118 bp and 2995 bp[298]. Polymorphism of FUT2 and its association with COVID-19 has been the subject of some studies. A genome-wide association was reported between rs368565 and critical COVID-19[278]. The rs368565 polymorphism may genetically pre-dispose an individual to develop severe COVID-9 illness.

1.3.3.14 HLA-G

As outlined on HLA play an important role in regulating the immune system. Pairo-Castineira *et al* identified a genome-wide association between rs9380142 HLA-G and critical COVID-19 illness ($p = 3.2 \times 10^{-8}$, OR = 1.3)[288]. Whereas Kousathanas *et al* reported genome-wide association between rs9271609 HLA-DRB1 and critical COVID-19 ($p = 3.26 \times 10^{-9}$, OR = 1.1). These results suggest that polymorphisms of HLA may confer susceptibility and severity of COVID-19.

1.3.3.15 IFITM3

Interferon-induced transmembrane protein (IFITM3) has an anti-viral role and helps prevent viruses from crossing the lipid bilayer[299]. IFITM3 on endocytic vesicles have been shown to fuse with virus particles and stimulate viral trafficking to lysosomes where they are degraded[300]. IFITM3 has also been shown to inhibit the entry of SARS-CoV-2[301]. Several polymorphisms of IFITM3 have been identified including rs12252 SNP which is located within the first exon of IFITM3 gene[302]. The rs12252 SNP, which results in the substitution of the majority T allele for C allele, modifies the splice acceptor site[303]. This results in a truncated form of the IFITM3 protein with a reduced antiviral capacity[304]. Mulla *et al* reported the C/C genotype of rs12252 was significantly associated ($p < 0.001$) with COVID-19 severity in a study population of 84 COVID-19 patients and 28 control subjects from Bangladesh[305]. The truncated form resulting from the rs12252 SNP could be a possible explanation for the disease severity however studies with larger population samples and structural-functional studies should be conducted to verify the role of the SNP. Whereas Schönfelder *et al* found no association between C allele or T/C genotype of rs12252 and A allele or G/A genotype of rs34481144 SNP and risk of SARS-CoV-2 severity in their analysis of 239 SARS-CoV-2 individuals and 253 control subjects[306]. Found on the regulatory region of IFITM3 gene, the rs34471144 SNP alters the transcriptional regulation by altering the binding affinity of transcription factors[306]. The deleterious CC genotype of the rs12252 SNP was only found in 2 SARS-CoV-2 positive individual[306]. Further research should be conducted using a larger cohort to validate the findings.

1.3.3.16 IFNAR2

Type 1 interferons (IFN-Is) are a cytokine family that are important for modulating the innate and the adaptive immune system against pathogens[307]. IFN-Is bind to cell surface receptors consisting of IFNAR1 and IFNAR2 subunits. IFNAR1 is associated with tyrosine kinase 2 (TYK2) whereas IFNAR2 is associated with Janus kinase 1 (JAK1)[307]. Upon IFN-1 binding, the two subunits are brought close together which drives cross-phosphorylation of TYK2 and JAK1 and the subsequent phosphorylation of tyrosine residues in the intracellular domain of the receptor subunits[307]. In individuals with severe

COVID-19, the levels of IFNAR2 were found to be higher than those with milder infection[308]. IFNAR2 polymorphisms and its association with COVID-19 have been the subject of several studies. Niemi et al reported association between rs13050728 IFNAR2 and SARS-CoV-2 infection[265]. Whereas Kousathanas *et al* identified an association between rs17860115 IFNAR2 and critical COVID-19 ($p = 9.69 \times 10^{-22}$, OR = 1.2)[278]. Moreover, Pairo-Castineira *et al* found a genome-wide association between rs2236757 and critical COVID-19 illness ($p = 5 \times 10^{-8}$, OR = 1.3)[288]. These data suggest that IFNAR2 polymorphisms may make an individual more susceptible to COVID-19 infection or/and develop a serious disease.

1.3.3.17 IL10RB

Intact interleukin-10 is important in the maintenance of immune tolerance. The IL10RB gene encodes two of the IL10Rβ molecules that form the heterotetrametric complex[309]. The remaining part of the complex, two IL10Rα, is encoded by IL10RA gene[309]. Upon binding of IL10 to IL10Rα, the sequence to assemble the receptor is initiated[309]. This activates TYK2 and JAK1 which is coupled to IL10Rβ and IL10Rα respectively resulting in phosphorylation of tyrosine residues found on the intracellular domain of IL10Rα[309]. Signal transducer and activator of transcription 3 (STAT3) is phosphorylated by TYK2 and JAK1 and translocates into the nucleus and then binds to the STAT3-binding elements of the IL10-responsive genes[309]. Polymorphism of IL10RB has also been associated with SARS-CoV-2 infection. GWAS study by Kousathanas *et al* found a genome-wide association between rs8178521 and critical COVID-19 ($p = 3.53 \times 10^{-12}$, OR = 1.2)[278].

1.3.3.18 IL6

Severe SARS-CoV-2 causes hyperactive host immune response with elevated pro-inflammatory cytokines including IL6[310]. Interference of tissue perfusion by IL6 and the subsequent endothelial damage worsens COVID-19 symptoms[311]. The association of IL6 rs1800795 and rs1800797polymorphisms with COVID-19 severity was investigated by Verma *et al*. The cohort of 97 patients with severe COVID-19 and 145 patients with mild COVID-19[312]. The G/C genotype of rs1800795 was associated with higher risk of disease severity (p = 0.009,

OR= 2.26)[312]. However, the rs1800797 polymorphism was not associated with disease severity[312].

1.3.3.19 KANSL1

KAT8 regulatory NSL complex subunit 1 (KANSL1), encoded by the KANSL1 gene, is a protein implicated in a rare genetic disorder called Koolen-de Vries syndrome[313]. The disorder is characterised by developmental delay, intellectual disability, and hypotonia[313]. Few studies have found an association between KANSL1 polymorphism and SARS-CoV-2 infection. A genome-wide association was identified between critical COVID-19 and rs2532300 polymorphism of KANSL1 ($p = 4.19 \times 10^{-9}$, OR = 1.2)[278]. Furthermore, Niemi et al reported a genome-wide association between rs1819040 and SARS-CoV-2 infection[265].

1.3.3.20 LINC00649

Long-coding RNA LINC00649 has been shown to mediate cellular process of bladder cancer cells by binding to microRNA-16-5p and increasing JARID2 expression[314]. LINC00649 has also been reported to target microRNA-216a-30 which may aggravate progression of cervical cancer[315]. Kousathanas *et al* reported a genome-wide association between rs35370143 and critical COVID-19 ($p = 1.24 \times 10^{-9}$)[278].

1.3.3.21 LZTFL1

Leucine zipper transcription factor-Like 1 (LZTFL1) is a protein associated with Bardet-Biedl syndrome (BBS)[316]. BBS is a rare genetic ciliopathy which affects multiple ciliated organs[316]. Moreover, LZTFL1 is shown to be important in sperm function and male fertility[316]. The LZTFL1 gene is located at chromosome 3 (3p21.3)[317]. Several studies have identified association between SARS-CoV-2 infection and LZTFL1 polymorphisms. The rs73064425 polymorphism has been associated with critical COVID-19 ($p = 1.97 \times 10^{-133}$, OR = 2.7) in a study by Kousathanas *et* al[278]. Pairo-Castineira *et al* also reported genome-wide association with critical COVID-19 infection ($p = 4.8 \times 10^{-30}$, OR = 2.1)[288]. Niemi *et al* identified genome-wide significance between rs10490770 and increased risk of developing COVID-19 symptoms ($p = 1.44 \times 10^{-73}$)[265].

1.3.3.22 MUC5B

MUC5B has an important function in mucociliary clearance and protection of the lung airways[318]. The gene for MUC5B lies on chromosome 11[319]. The rs35705950 polymorphism is implicated in MUC5B expression in the lungs[319]. The polymorphism could potentially be involved in the methylation of the promoter region of MUC5B and disrupt the activities of transcriptional factors[319]. Van Moorsel *et al* conducted a case-control study to investigate the influence of this polymorphism on COVID-19. The study population consisted of 83 white Dutch patients with COVID-19 and 611 white Dutch healthy individuals as control subjects[320]. The frequency of the T allele was significantly lower in the COVID-19 group compared to controls ($p = 0.02$)[320]. This suggests that the T allele may provide protection to individuals from developing a severe disease.

1.3.3.23 NOTCH4

Neurogenic locus notch homolog 4 (NOTCH4) is a protein from the NOTCH family. NOTCH signalling is an evolutionary conserved pathway that is implicated in various developmental processes as well as several cancers[321]. Expression of NOTCH4 in T regulatory cells has been shown to impair tissue repair to stimulate severe lung inflammation[322]. The gene encoding NOTCH4 is localised on a small arm of chromosome 6 (6p21.3)[323]. Polymorphism of NOTCH4 and its influence on COVID-19 has been reported in literature. A genome-wide association between rs3131294 and critical COVID-19 illness was identified by Pairo-Castineira ($p = 2.8 \times 10^{-8}$, OR = 1.5)[288]. Furthermore, in individuals with COVID-19, the expression of NOTCH4 is upregulated on circulating T regulatory cells and associated with amplified mortality[322].

1.3.3.24 OAS1-3

As sensors of the innate immune system, 2',5'-oligoadenylate synthetase (OAS) binds to viral dsRNA and initiates degradation via ribonuclease L to reduce the spread of viral infection[324]. The gene cluster on chromosome 12 (12q24.13) encodes for OAS1, OAS2 and OAS3 genes[324]. The influence of OAS1-3 polymorphisms in COVID-19 infection have been investigated in several studies. GWAS study by Pairo-Castineira *et al* found genome-wide

significance association in risk allele A of rs10735079 of the OAS gene cluster ($p = 1.6 \times 10^{-8}$, OR = 1.3)[288]. Association between rs10774671 and increased risk of developing severe COVID-19 symptoms was reported by Niemi *et al* ($p = 6.14 \times 10^{-10}$)[265]. This suggests that anti-viral defences of the immune system are associated with COVID-19 severity.

1.3.3.25 PLSCR1

Phospholipid scramblase 1 (PLSCR1) is a type II transmembrane protein with a role in phospholipid movement across the plasma membrane[325]. PLSCR1 is also able to modulate phagocytic processes in differentiated macrophages[325]. The PLSCR1 gene is localised on chromosome 3 (3q24) and comprises nine exons with the coding region spanning from exon 2 to exon 9[326]. Polymorphism of PLSCR1 has been reported to be associated with severity of COVID-19. A genome-wide association between rs343320 and critical COVID-19 was reported by Kousathanas *et al*[278].

1.3.3.26 RGMA

Repulsive guidance molecule A (RGMA) is a Bone Morphogenetic Protein (BMP) co-receptor that facilitates BMP signalling and activates SMAD1/5/8 signalling pathway[327]. The RGM loci which encodes RGMA is mapped onto chromosome 5 (15q26.1). The rs4424872 polymorphism of the RGMA gene showed genome-wide association with critical COVID-19 ($p = 8.61 \times 10^{-13}$, OR = 2.4)[278]. This suggests that the polymorphism may increase the risk of developing critical COVID-19 illness.

1.3.3.27 Solute Carrier Family and Genes of Chromosome 3

Solute carrier family (SLC) is a superfamily which comprises 65 families and 458 proteins[328]. SLCs respond to different metabolic states by transporting substances across the cellular membranes[328]. SLC6 family carry out active transport of small amino acids using the sodium gradient[328]. One of the members of SLC6, Solute carrier family 6 member 20 (SLC6A20), is a sodium chloride-dependent transporter and a member of the neutral amino acid transporter subfamily. The SLC6A20 is found on chromosome 3p21.3. Kousathanas identified a genome-wide association between rs2271616 and

critical COVID-19 ($p = 9.9 \times 10^{-17}$). Niemi *et al* also reported an association with rs2271616 polymorphism. The polymorphism was associated with increased susceptibility to SARS-CoV-2 infection ($p = 1.79 \times 10^{-34}$)[265]. Whereas the Severe COVID-19 GWAS Group found that the GA or G allele of rs11385942 was associated with COVID-19 ($p = 1.15 \times 10^{-10}$, OR = 1.77). Six genes comprised the rs117169628 association signal at locus 3p21.31: SLC6A20, LZTFL1, C-C chemokine receptor type 9 (CCR9) CCR9, FYVE And coiled-coil domain containing 1 (FYCO1), C-X-C chemokine receptor type 6 (CXCR6), and XCR1. As an inflammatory and homeostatic receptor, CCR9 has been shown to be an impotent mediator of the early phases of allergic airway inflammation[329]. FYCO1 is an adapter protein with roles in transportation of autophagic vesicles and is implicated in the development of cataracts[330]. CXCR6 is a receptor for simian immunodeficiency virus[331]. XCR1 is the receptor for XCL1 and the axis comprising the two is key in the dendritic-cell-mediated cytotoxic immune response. Moreover, Shelton *et al* reported an association between critical COVID-19 and rs13078854 polymorphism[332]. Another SLC reported to be associated with COVID-19 is solute carrier family member 31 (SLC22A31). The rs117169628 polymorphism was associated with critical COVID-19 ($p = 4.4 \times 10^{-9}$).

1.3.3.28 THBS3

Thrombospondin 3 (THBS3) is a glycoprotein with an important role in interactions between cell-cell and cell-matrix[333]. THBS3 has been shown to be a suitable marker for acute myocardial infarction[333]. Located on chromosome 1, polymorphism of THBS3 has been reported to mediate a genetic susceptibility to critical COVID-19 illness. Kousathanas *et al* reported a genome-wide association between rs41264915 and critical COVID-19 (1.02×10^{-12}, OR = 1.3)[278].

1.3.3.29 TMEM64

Transmembrane protein 65 (TMEM65) is found within the inner mitochondrial membrane[334]. Although not much is known about its function, introduction of a TMEM gene mutation hampered the content of the mitochondria and the respiration rate[334]. The rs72711165 polymorphism was associated with increased risk of developing severe COVID-19 symptoms ($p = 2.13 \times 10^{-9}$)[265].

1.3.3.30 TYK2

TYK2 is a member of the JAK family and is an important component of the JAK-STAT pathway[335]. TYK2 helps regulate cytokine signal transduction pathways and immune cell functions[336]. Several studies have reported an association with TYK2 polymorphism and COVID-19. Genome-wide significant association was reported in allele A of rs74956615 near a gene that encodes for TYK2 on chromosome 19 (19p13.3) ($p = 2.3 \times 10^{-8}$, OR = 1.6)[288]. High expression of TYK2 was also related with life-threatening disease[288]. Niemi *et al* also reported genome-wide association between rs74956615 and increased risk of developing severe COVID-19 symptoms ($p = 5.05 \times 10^{-10}$)[265]. Whereas Kousathanas *et al* also reported genome-wide association between rs34536443 and critical COVID-19 ($p = 6.98 \times 10^{-17}$, OR = 1.5)[278]. Furthermore, expression of TYK2 has shown to be downregulated in patients with COVID-19[337]. This could potentially hinder the ability to produce a competent IFN response upon SARS-CoV-2 infection[337].

1.3.4 Summary

Several host genetic factors have shown association with pre-disposition and disease outcome of RV, RSV and SARS-CoV-2 infection. These include a diverse repertoire of host factors including components of the cytokine pathway, transmembrane receptors, and surfactant proteins. The differences in allelic frequencies of polymorphisms could be a possible explanation for the variation of COVID-19 severity observed in different population groups as well as the difference in risk allele predisposing to RV and RSV infection in different ethnic groups. Genetic analysis of polymorphisms should be conducted using a wide range of different population groups for comparison. Diverse range of experimental designs have been utilised when investigating these polymorphisms such as large-scale GWAS studies and case-control studies. The use of case-control analysis of polymorphisms in viral infections allows us to visualise the influence of the polymorphism on clinical outcome. However, the polymorphisms could possibly be a marker for other uncharacterised polymorphisms and so it can be difficult to decipher whether the association is a result of a specific allele. Nevertheless, studies using a large cohort such as GWAS studies will give a better understanding on whether an allele or the

genotype of a particular polymorphism is associated with disease susceptibility and severity. Assessing the influence of polymorphisms on the susceptibility and severity to the different viral infection is imperative for understanding the mechanisms underlying the pathogenesis. Furthermore, this will help enhance drug development and prophylaxis against viral infections. In addition, the polymorphisms may serve as biomarkers and so can be used for targeted interventions in high-risk individuals.

1.4 Modelling the Airway Epithelium to Study Host-Virus Interactions

Modelling the airway epithelium *in vitro* is an important tool for understanding how respiratory viruses interact with the host. The bronchial airway epithelium is typically modelled using primary cells or continuous cell lines. Animal models do not faithfully represent the human respiratory epithelium whereas *ex vivo* human lung samples have issues with reliability and availability[338]. Therefore cell culture models enable closer recapitulation of the airway epithelium. By studying host-virus interactions using cell culture models, advances in disease modelling and development of therapeutics can be made.

1.4.1 Primary Cell Culture

Culturing of primary cells via air-liquid interface (ALI) is the current gold standard for modelling host-respiratory virus interaction. Human bronchial epithelial cells (HBECs) are obtained from non-diseased lungs to exhibit the 'normal' pulmonary phenotype. The bronchoscopy brush technique is applied where the epithelial cells from the bronchial mucosa are scraped using a brush. The HBECs are cultured in submerged conditions and then transferred to collagen-coated porous membrane inserts following confluency. The cells are exposed to air on the apical side and to the cell culture media at the basolateral side[339]. This enables mucociliary differentiation, tight junction formation and development of specialised cells such as ciliated cells and mucus-producing goblet cells. This allows the pseudostratified morphology of the *in vivo* bronchial airway epithelium to be recapitulated[340]. Since the architecture of the airway epithelium resembles the *in vivo* structure, HBEC-ALI model is one of the most routinely used culture techniques for studying host-virus interactions[340]. Figure 1-9 shows the difference in morphology of ALI cultures at different time points[341].

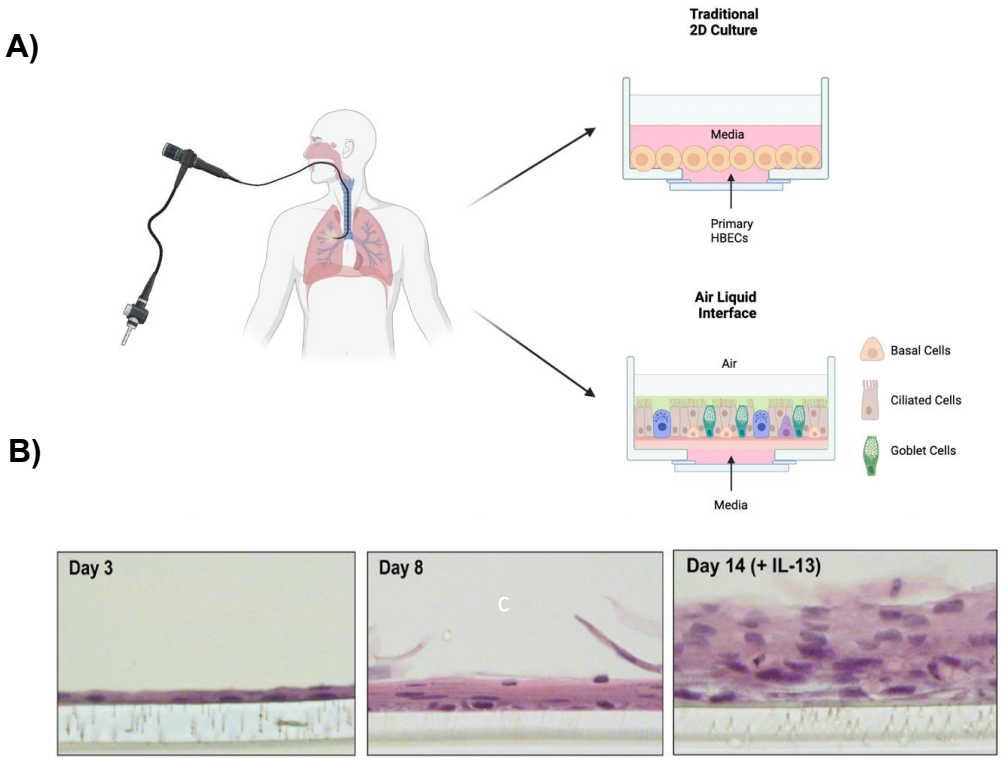


Figure 1-9. Traditional 2D Culture vs ALI Culture. A. Cells are taken from bronchoscope and cultured in submerged in conditions using cell culture media. Cells cultured in air-liquid interface have the ability to form a pseudostratified layer with specialised cells such as basal, ciliated and goblet cells. Diagram was created using BioRender with Kiran *et al* used as reference[15]. **B.** Images of air-liquid interface culture taken at different time points. The cells are stained using haemotoxylin and eosin[15].

1.4.2 Continuous Cell Lines

Continuous cell lines are typically cultured as a 2D monolayer culture in submerged conditions. One example, the lung adenocarcinoma-derived Calu-3 cell line, exhibits many characteristics of the bronchial airway epithelium[338]. They can form tight polarised monolayers, and can express mucin, surfactant protein C and CFTR[342]. Another cell line, BEAS-2B, are routinely used in infection studies. The cell line is established from normal human bronchial epithelial from a non-cancerous individual[343]. Although BEAS-2B express E-cadherin, ICAM-1, CCSP, and vascular cell adhesion molecule (VCAM-1), the

cells do not demonstrate the typical cellular distribution of the bronchial airway epithelium[338]. The caveat of several continuous cell lines being cancerous makes them less than ideal for representing the bronchial airway epithelium. In addition, the inability to form specialised cells makes the model less physiological relevant than using primary cells.

1.4.3 Induced Pluripotent Stem Cells

The development of induced pluripotent stem cells (iPSC) has transformed *in vitro* modelling of human biology and therapeutic development. Martin Evans and Mathew Kaufman[344], and Gail Martin[345] first isolated mouse embryonic stem cells (ESCs) in 1981 which would serve as a reference point for subsequent experiments involving somatic cell reprogramming. James Thomson *et al* later isolated human ESCs[346]. Cell fusion experiments of human ESCs with somatic cells showed the capability of reprogramming cells to pluripotency[347]. Whereas transdifferentiation experiments highlighted the importance of transcription factors in orchestrating the cell fate[348]. In 2006, Takahashi and Yamanaka conducted a series of somatic cell reprogramming experiments which led to the breakthrough generation of mouse iPSCs[349]. A selection of 24 potential reprogramming factors was cloned into retroviral vectors for mouse embryonic fibroblasts (MEFs) transduction[349]. Following screening, four transcription factors (known as OSKM or Yamanaka factors) were found to be sufficient to stimulate pluripotency: octamer binding transcription factor (OCT3/4), SRY-box 2 (SOX2), c-Myc, Krupper-like factor 4 (KLF4)[349]. During the early stage of reprogramming, OSKM expels somatic cell-specific transcription factors from somatic enhancers[350]. In particular they access closed chromatin and initiate gene expression[351]. OCT4, SOX2 and KFF4 have a pioneer activity where they target partial motifs in the nucleosome-enriched loci[351]. In addition, SOX2 promotes DNA demethylation and impedes the repressive epigenome[352]. Pluripotency enhancers are then activated by OSKM during later stages of reprogramming[350]. Yamanaka and Takahashi also demonstrated that human fibroblasts can be reprogrammed into iPSCs[353].

The capacity for differentiating iPSCs into a diverse repertoire of human cell types and reduced ethical constraints has enabled iPSCs-derived cells to be widely used in disease modelling, high-throughput drug screening, and developing therapies[354]. The inherent property of self-renewal allows large quantities of cells to be generated[355]. This overcomes the limitation of finite lifespan of primary cells[355]. Another important advantage of iPSCs is that cells can be derived from patients with genetic risk factors or casual disease mutation, or a particular disease which allows patient-specific cell lines with relevant genetic background to be generated and enables robust modelling and identification of disease-specific phenotypes[354]. Furthermore coupling iPSCs with gene-editing technologies such as CRISPR/Cas9 allows mutations to be introduced and studied. This provides opportunity to investigate further into host genetic factors and its influences including on host-virus interactions[355].

1.5 Using Transcriptomics to Understand Host-Virus Interactions

The transcriptome is the complete set of coding and non-coding (ncRNA) RNA transcripts in a cell of an organism[356]. mRNA are frequently studied since they encode proteins via the genetic code[357]. However, there is a diverse group of functional non-coding ncRNA. Examples include ribosomal RNAs role in mRNA translation, small nuclear RNA in splicing, small nucleolar RNAs in rRNA modification, and microRNA in gene expression regulation at post-transcriptional level[357]. Genomic research has been revolutionised by the use of transcriptomic technologies[358]. Thousands of gene transcripts can be simultaneously examined by using a genome-wide transcriptome profiling[359]. This allows functional elements of the genome and molecular constituents of cells and tissues to be interpreted[356]. Several techniques have been developed to elucidate the transcriptome. One such example, the DNA microarray, is based on a collection of spots that is occupied by a fragment of single-stranded DNA sequence known as probes[360]. The DNA sequence hybridises with its complementary strand[360]. This target-probe hybridisation is identified and quantified by detecting fluorophore or chemiluminescence-labelled targets[360]. However limitations of microarrays include the availability of probes for known genes, prior knowledge of sequences being interrogated, cross-hybridisation artifacts, low abundance transcripts, and limited ability to accurately quantify very highly expressed and lowly expressed genes[359].

Sequence-based approach, where the transcript sequence is directly determined, was developed in response to issues with microarray-based methods[357]. Tag-based Sanger sequencing methods such as serial analysis of gene expression (SAGE) and cap analysis gene expression (CAGE), quantifies the number of tagged sequences[356]. This directly corresponds to the number of mRNA transcripts and so allows a more precise quantification of expression levels than microarray methods[357]. However, they cannot be used for novel gene discovery and are also insensitive for measuring the levels of splice isoforms expression[357]. Furthermore, the need for large volumes of RNA and the high cost of Sanger sequencing has restricted its use[357]. Transcriptomics transitioned to sequencing-based methods following issues

with microarray and tag-based methods. RNA-sequencing (RNA-Seq) is a high-throughput DNA sequencing method that maps and quantifies the transcriptome[356]. It typically involves isolation of RNA, conversion of RNA to complementary DNA (cDNA), preparation of sequencing library, and sequencing using next-generation sequencing (NGS) platform[357]. The dynamic nature of the transcriptome can be explored at greater resolution and higher coverage with RNA-seq compared to microarray and tag-based methods. Furthermore, recent advances in the workflow with sample preparation, sequencing platform, and data analysis has facilitated deeper profiling and opportunity to examine different pathological and physiological conditions[357]. In comparison to microarrays, RNA-seq is able to detect novel splice forms, closely related gene sequences, and RNA editing that may have potentially been missed due to cross hybridisation[361]. Furthermore, RNA-seq can be used to reveal the precise location of transcription boundaries to a single-base resolution[356]. A 30-bp short reads can provide information about how two exons are connected. In addition, a short pair-end reads or longer reads can unveil connectivity between multiple exons[356]. RNA-Seq can also be used to identify SNPs in transcribed regions[356].

Transcriptomic profiling using RNA-seq provides opportunities to investigate host-virus interactions at high-resolution. Viral infection triggers a dynamic set of reactions in the host and subsequently leads to changes in patterns of gene expression[362]. This can trigger the host immune response to eliminate the virus or to confer adaptation and tolerance towards the virus[362]. The complex interactions between the host and virus can influence the life cycle of the virus, the ability to infect, and the ability to evade the host immune system[363]. These complexities can affect disease severity[363]. Understanding these interactions is imperative for developing new therapeutics since viruses rely on host biological mechanisms for replication and propagation[363]. By interrogating the transcriptome using RNA-Seq, the transcriptional dynamics of the host and virus during infection can be explored[364]. The changes in the abundance of transcripts across samples can be detected to identify genes and pathways that are upregulated or downregulated in response to infection[364].

Furthermore, virulence factors or PAMPS targeting the viruses can be identified therefore deepening our understanding of host-virus interactions[362].

1.6 Aims and Objectives

1.6.1 Overall Aim

To understand host-virus interactions at genomic and transcriptomic level and investigate the interactions further using HBEC and iPSC model of the airway epithelium.

1.6.2 Objectives

1. To explore host-virus interactions using genomic-based approach by conducting a signal-to-gene analysis to translate genetic signals associated with RV, RSV and SARS-CoV-2 susceptibility and severity and conducting a pathway, diseases and drug-gene interaction analysis.
2. To further understand host-virus interactions using transcriptomics by investigating how primary human bronchial epithelial cells respond to prolonged infection of RSV-A2 and RV-A16.
3. To develop iPSC model of the airway epithelium and compare and contrast with the gold standard ALI-HBEC model of the airway epithelium to investigate host-RSV interactions.

2.0 Materials and Methods

2.1 Molecular Biology Methods

This section details the molecular biology methods used during the thesis. Details of RNA-seq and reverse transcription quantitative polymerase chain reaction (RT-qPCR) are outlined in Chapter 4 and Chapter 5 respectively.

2.1.1 Material and Reagents

The table 2-1 below details the material and reagents used for molecular biology methods.

Table 2-1. List of materials and reagents used for molecular biology methods. The supplier and the catalogue code are stated on the table.

Materials and Reagents	Supplier
RNeasy Mini Kit (cat. no. 74104)	Qiagen
QIAshredder (cat. no. 79654)	Qiagen
RNase-free DNase Set (cat. no. 79254)	Qiagen
Ethanol (cat. no. LH-M0120E)	Medical School Stores
TaqMan™ Gene Expression Master Mix (cat. no. 4369016)	ThermoFisher Scientific
TaqMan® Gene Expression Assays FOXA2 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays SOX17 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays NKX2-1 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays FOXJ1 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays KRT5 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays MUC5AC (cat. no. 4331182)	ThermoFisher Scientific

TaqMan® Gene Expression Assays RSV A2 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays CX3CR1 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays OAS1 (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays IFNB (cat. no. 4331182)	ThermoFisher Scientific
TaqMan® Gene Expression Assays IFNL (cat. no. 4331182)	ThermoFisher Scientific

2.1.2 RNA Extraction

RNA extraction is a routinely molecular biology technique to isolate RNA for downstream applications such as quantitative polymerase chain reaction (qPCR) and RNA-seq[365]. The cell lysates are loaded onto the silica membrane which enables binding of RNA to the silica membrane. This ideal binding condition is facilitated by the addition of ethanol. High quality of RNA is important for ensuring accurate analysis of gene expression[365]. Therefore, DNA digestion was conducted for a complete DNA removal.

Protocol

1. Cells were lysed using 350 µl of Buffer RLT and scraped using a pipette
2. To make the DNA incubation mix, 10 µl of DNase I stock solution was added to 70 µl of Buffer RDD
3. The tube containing the DNA incubation mix was gently inverted and briefly centrifuged to collect residual liquid from the sides of the tube
4. Cell lysate was transferred into QIAshredder spin column to homogenise the cells and prevent reduction in RNA yields and clogging of the RNeasy spin column
5. The QIAshredder spin columns were centrifuged at ≥12000 rpm for 2 minutes
6. 350 µl of 70% ethanol was added to the homogenised lysate and mixed thoroughly using a pipette
7. The sample was transferred to a RNeasy Mini spin column which was connected to a 2 ml collection tube
8. The spin column was centrifuged at ≥8000 *g* (≥10000 rpm) for 15 seconds
9. The sample was incubated with 80 µl of DNA incubation mix for 15 minutes at 20-30°C
10. 350 µl of Buffer RW1 was added to the spin column
11. The spin column was centrifuged at ≥10000 rpm for 15 seconds and the flow-through being was discarded
12. 500 µl Buffer RPE was added to the RNeasy spin column
13. The spin column was centrifuged at ≥10000 rpm for 15 seconds and the flow-through was discarded

14. 500 µl Buffer RPE was added to the RNeasy spin column
15. The spin column was centrifuged at ≥10000 rpm for 15 seconds
16. The old collection tube with the flow-through was discarded and the spin column was placed in a new 2 ml collection tube and centrifuged for 1 minute at ≥12000 rpm
17. The RNeasy spin column was then placed in a new 1.5 ml collection tube
18. 30 µl of RNAse-free water was added to the spin column membrane and centrifuged at ≥10000 rpm for 1 minute to elute the RNA
19. The RNA samples were stored at -80°C

2.1.3 Quantification of RNA

The NanoDrop ND-2000 full spectrum spectrophotometer (ThermoFisher Scientific) was used to quantify RNA concentration. The purity of RNA with respect to protein contamination was assessed using the 260/280 ratio. The measurement of purity with regards to contamination from reagents such as salts is reflected by the 260/230 ratio.

Protocol

1. 1µl of RNase-free water was placed at the lower measurement pedestal and the sampling arm was lowered
2. Spectral measurement was initiated using the NanoDrop 2000/2000c PC software to blank the sample
3. 1µl of the RNA sample was added to the lower measurement pedestal to measure the concentration

2.2 Cell Culture Methods

Cell culture process comprises growing cells under a set of controlled conditions involving temperature, pressure and gases[366]. The cells are typically grown at 37°C and 5% CO_2 and cultured in medium optimum to their growth. Cells grow in a monolayer following adherence to a surface. The cells are passaged once they reach confluence, and a portion of the cells are then cultured in a new flask. Cell culture can involve primary cultures whereby *ex vivo* culture of cells is obtained from human tissues, or immortalised cell lines[366]. Although primary cell cultures are more physiologically relevant, they have a limited life-span[366]. In contrast, immortalised cells lines possess a longer lifespan[366]. However they lack physiological relevance due to lower expression of important physiological functions and genetic alteration[366]. This section outlines primary and continuous cell lines used during the project. Methods outlining iPSC culture are detailed in Chapter 5.2.

2.2.1 Materials and Reagents

The following section details the cell lines and reagents used for cell culture for this Chapter. Table 2-2 lists the cell culture reagents that were used.

Table 2-2. Specifications for HBEC donor D485960 used for cell culture experiments.

Reagent	Supplier
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM F-12) (cat. no. 11554546)	ThermoFisher Scientific
Dulbecco's Phosphate Buffered Saline (PBS) (cat. no. D8537)	Sigma-Aldrich
Gibco™ TrypLE™ Select Enzyme (1X) (cat. no. 11598846)	ThermoFisher Scientific
KnockOut™ DMEM (cat. no. 10829018)	ThermoFisher Scientific
KnockOut™ Serum Replacement (cat. no. 0828010)	ThermoFisher Scientific
Gibco™ Penicillin-Streptomycin (10,000 U/mL) (cat. no.)	ThermoFisher Scientific
GlutaMAX™ Supplement (cat. no. 35050061)	ThermoFisher Scientific
1-Thioglycerol (cat. no. M6145)	Sigma-Aldrich
Gibco™ MEM Non-Essential Amino Acids Solution (100X) (cat. no. 11140050)	ThermoFisher Scientific
Human Collagen Type IV (C7521)	Sigma-Aldrich

Animal-Free Recombinant Human FGF-basic (cat. no. AF-100-18B)	PeproTech
Recombinant Human Sonic Hedgehog/Shh (C24II) N-Terminus (cat. no. 845-SH-025)	R&D Biosystems
Recombinant Human KGF/FGF-7 Protein (cat. no. 251-KG-010)	R&D Biosystems
Recombinant Human FGF-18 (cat. no. 100-28)	PeproTech
Collagen from Human Placenta (cat. no. C7521)	Sigma-Aldrich
B-ALI™ Bronchial Epithelial Cell Growth Medium BulletKit™ (cat. no. CC-3170)	Lonza
B-ALI™ Bronchial Air-Liquid Interface Medium BulletKit™ (cat. no. 00193514)	Lonza
PureCol® Bovine Collagen (cat. no. 5005)	Advanced BioMatrix
Gibco™ Trypsin-EDTA Trypsin (cat. no. 10779413)	ThermoFisher Scientific
Soybean Trypsin Inhibitor (cat. no. 17075029)	
DMEM, high glucose, GlutaMAX™ Supplement, pyruvate (cat. no. A5256701)	ThermoFisher Scientific
Fetal Bovine Serum (cat. no. A5256701)	ThermoFisher Scientific

PneumaCult™-Ex Plus Medium (cat. no. 05040)	STEMCELL Technologies
Hydrocortisone Stock Solution (cat. no. 07926)	STEMCELL Technologies
PneumaCult™-ALI Medium (cat. no. 05021)	STEMCELL Technologies
DMSO (cat. no 154938)	Sigma-Aldrich
NHBE-Bronchial Epithelial Cells D485960 (cat. no.CC-2540S)	Lonza
NHBE-Bronchial Epithelial Cells D195307 (cat. no.CC-2540S)	Lonza
iPSC LOPCK	Dr Nicholas Hannan
iPSC ReBL-PAT	Dr Nicholas Hannan
HRVA16 (cat. no. 0112092v)	Public Health England
RSV-A2	Professor Colin Bingle
RSV-A2 GFP	Professor Colin Bingle

2.2.2 HBECs

The methods for maintaining HBECs in submerged condition are detailed in this section. Details on culturing HBECs in ALI conditions are outlined in Chapter 5. HBEC donors D485960 and D195307 were purchased from Lonza. The specification of the donors is detailed on Table 2-3 and 2-4.

Table 2-3. Specifications for HBEC donor D485960 used for cell culture experiments.

Test	Minimum	Maximum	Results
Tissue Acquisition Number			28968
Donor Screen Information			
Age			38 Y
Sex			MALE
Race			CAUCASIAN
Cell Type			NHBE
QC Evaluation Medium			BEGM
Cell Strain Calculations			
Cell Passage			1
Viability	>= 70%		83%
Cell Count (cells/amp)	>= 500,000		540,500
Seeding Efficiency	>= 20%		45%
Doubling Time (hours)	12	48	35
Sterility Test			Negative
Mycoplasms			Negative
Virus Testing			
HIV Test			Not Detected
HBV Test			Not Detected
HCV Test			Not Detected

Table 2-4. Specifications for HBEC donor D195307 used for cell culture experiments.

Test	Minimum	Maximum	Results
Tissue Acquisition Number			19665
Donor Screen Information			
Age			19 Y
Sex			MALE
Race			CAUCASIAN
Cell Type			NHBE
QC Evaluation Medium			BEGM
Cell Strain Calculations			
Cell Passage			1
Viability-Tryp Blue Exclusion	>= 70%		91%
Cell Count (cells/amp)	>= 500,000		897,000
Total Population Doublings	>= 15		15
Seeding Efficiency	>= 20%		66%
Doubling Time (hours)	12	48	15
Sterility Test			Negative
Mycoplasms			Negative
Virus Testing			
HIV Test			Not Detected
HBV Test			Not Detected
HCV Test			Not Detected

2.2.2.1 Coating of Cell Culture Dishes for HBECs

To allow attachment of HBECs to the surface, cell culture dished were coated with PureCol® Bovine Collagen. As a widely used extracellular membrane component coating in cell culture, collagen enables adhesion to cell culture vessels[367].

Protocol

1. A collagen-phosphate-buffered saline (PBS) solution was prepared in a 1:100 ratio of collagen and PBS solution
2. 10 ml of the collagen solution was added to T25 flask
3. T25 flask was incubated for at least 30 minutes 37°C and 5% CO_2 in an incubator

2.2.2.2 Maintenance of Submerged HBECs

Protocol

1. Vial of HBECs frozen in liquid nitrogen stocks were thawed and seeded onto T75 flask
2. 11 ml of Lonza BEGM® Bronchial epithelial cell growth medium bulletKit® was added to the T75 flask
3. The cells were incubated at 37°C with 5% CO_2 in a cell culture incubator
4. Media was replaced every 48 hours and replaced last thing on Friday afternoon and first thing on Monday morning

2.2.2.3 Passaging of Human Bronchial Epithelial Cells

Protocol

1. Upon reaching confluency, 2.5 ml of trypsin (2 mg/ml)/ ethylenediaminetetraacetic acid (EDTA) (5 mg/ml) was added to flask containing the HBECs and incubated for 5 minutes at 37°C
2. The cells were checked for detachment using the microscope
3. 7.5 ml of fresh PneumaCult™ BEGM media was added to the T75 flask and pipetted up and down to break up the clumps
4. 5 ml of trypsin inhibitor was added
5. The cell suspension was transferred to a universal tube
6. The cells were centrifuged at 200 rpm for 5 minutes

7. The cells were resuspended in 1 ml of Lonza BEGM® and replated into a new collagen coated-T25 flask with 11 ml of fresh media
8. The cells were incubated at 37°C with 5% CO_2 in a cell culture incubator
9. Cells were passaged 1:4 routinely

2.2.2.4 Freezing of Human Bronchial Epithelial Cells Protocol

1. HBECs were counted using the Cell Drop Machine
2. Freezing solution was prepared by adding PneumaCult ex plus media to sterile dimethyl sulfoxide (DMSO) in a 9:1 ratio
3. The cell suspension was diluted to 250,000 cells/ml using the freezing solution
4. The cell solution was quickly transferred into a labelled cryovial
5. The cryovials were placed inside Mr Frosty and stored in -80°C freezer for 48 hours
6. After 48 hours, the cryovials were transferred to liquid nitrogen

2.2.3 HEK293-T

The HEK293-T cell line was first established following transformation of human embryonic kidney cells with sheared adenovirus type 5 DNA[368]. A derivative of the HEK293, the HEK293-T was produced by enabling a temperature-sensitive SV40 T-antigen mutant to be expressed[368]. HEK293-T cells were used to propagate the viruses RSV-A2 and RVA-16.

2.2.4.1 Maintenance of HEK293-T

Protocol

1. T25 flask was coated with collagen as outlined on section 2.2.3.1
2. 10 ml of Dulbecco's modified eagle medium (DMEM) supplemented with 10% foetal calf serum (FCS) and 1% penicillin-streptomycin was added to the T25 flask
3. Vial of HEK293-T was thawed and seeded onto T75 flask
4. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator
5. Media was replaced every 48 hours

2.3 Viral Methods

This section details the viral methods used during the thesis: propagation, determination of MOI and infection of iPSCs and HBECs using RSV-A2 and RV-A16. RSV-A2 was kindly provided by Professor Colin Bingle[369]. RV-A16 was purchased from Public Health England (Catalogue code: 0112092v).

2.3.1 Propagation of RSV-A16 and RV-A16

Protocol

1. 6.3 x 10^6 HEK293-T cells in 68 ml of DMEM supplemented with 10% Foetal Bovine Serum and 1% penicillin-streptomycin were seeded onto T225 flask
2. The cells were incubated at 37ºC and 5% CO_2
3. Media was replaced every 48 hours until confluency
4. Upon confluency, the cell monolayer was washed twice using 45 ml
5. The cells were infected with 0.01 MOI of virus with a total volume of 200 µl
6. The cells were incubated at 37ºC and 5% CO_2 for 72 hours
7. After 72 hours, the media was harvested from the T225 flask using a 2ml glass pipette
8. The media was filtered using a 0.45 µm filter to remove the cellular debris
9. The viral supernatant was transferred into a universal
10. The universal containing the supernatant was sealed using a parafilm and placed on a roller in a cold room overnight at 4ºC
11. The viral supernatants were harvested at 24, 48 and 72 hours and pooled together
12. The pooled supernatants were aliquoted into centrifuge tubes and sealed using parafilm
13. The supernatants were ultracentrifuged at 20,000 rpm for 3 and a half hours at 4ºC
14. Following ultracentrifugation, supernatant was aspirated carefully
15. The pellet was resuspended without FCS
16. The resuspended pellet was incubated at 4ºC overnight on a roller
17. The virus was aliquoted in 200 µl aliquots and stored at -80ºC

2.3.2 Determination of Multiplicity of infection (MOI) for Respiratory Syncytial Virus and Rhinovirus

Protocol

1. HBECs were seeded onto 96-well plate with 200 µl of media and incubated at 37°C with 5% CO_2
2. Media was changed every 48 hours until the cells are confluent
3. An 8 point 10-fold serial dilution of the virus suspension including a negative control comprising of media alone was set up
4. Media was removed from the culture plate, and the virus dilution was added to designated wells in triplicate
5. Using a light microscope, cell cultures were monitored on a daily basis until a 50% endpoint (50% cells are alive, and 50% cells are dead in the well) was reached
6. The tissue culture infectious dose (TCID50) was calculated using the Reed-Muench formula

2.3.3 Infection of iPSCs and HBECs with RSV-A2 and RV-A16

Protocol

1. Cell culture media was aspirated
2. The cells were washed with warm PBS and aspirated
3. The cells were infected with RSV-A2 or RV-A16 with an MOI of 10
4. The cells were incubated at 37℃ and 5% CO_2 in a cell culture incubator
5. After 1 hour, the virus-media was removed
6. The cells were washed with warm PBS and fresh media without the virus was added
7. The cells were incubated at 37℃ and 5% CO_2 in a cell culture incubator

3.0. Translating Host-respiratory Virus Genetic Susceptibility and Severity Signals to Candidate Casual Genes and Pathways

3.1 Introduction

An accumulating number of studies suggest that host genetic factors influence the susceptibility and severity to respiratory viral infections. Association between single nucleotide polymorphisms (SNPs) and susceptibility and severity to viral infections have been identified in various genome-wide association studies and case-control candidate gene studies. GWAS tests millions of genetic variants to identify association between variants and traits[370]. This has led to successful identification of novel genotype-phenotype association which has facilitated identification of new drug targets, and biomarkers, disease-causing mechanisms and pathways, and optimisation of therapeutics based on genotype of the individual[370]. For example, the identification of rs2241880 polymorphism of autophagy related 16 Like 1 (ATG16LI) led to the establishment of the role of autophagy in Crohn's disease[371]. The genetic signal results in a missense mutation that promotes the caspase-3-mediated cleavage of ATG16L1 during cellular stress and suppresses autophagy[371]. Although initial identification of a locus can be ascertained, it remains difficult to accurately pinpoint causal variants and genes[370]. Many signals map onto non-coding regions and so may not be involved in the functional role of the gene[370]. Therefore, it is important to accurately determine the casual genes from the signals[370].

The number of genetic signals found to be associated with pathogenesis of rhinovirus infection are relatively low compared to RSV and SARS-CoV-2. This could be explained by the low number of studies focusing on the effect of genetic signals on RV infection. A study by Doyle *et al* found that low production IL6 conferred by C/C genotype of rs1800795 was associated with greater RV illness[372]. Another study reported that individuals with the rs1800890 and rs1800896 polymorphism of IL10, which is associated with elevated IL10 production, presented less with RV-induced bronchiolitis[146]. The rs1060826 NOS2A gene was also associated with bronchiolitis induced by RV

infections[373]. A great number of studies on polymorphisms associated with RSV susceptibility and severity have focused on cytokines such as IL4, IL6, IL8, IL10, IL15 and IL8. Zhang *et al* reported that the T allele of rs2243250 polymorphism of IL4 was significantly higher in the RSV group compared to control[186]. Similar to the study by Doyle *et al*, Gentile *et al* found that the low production IL6 rs1800795 polymorphism was associated with RSV disease severity[179]. Higher frequency of the A allele of rs4071 was found among subjects with RSV-induced bronchiolitis by Hull *et al* and Goethghebuer *et al*[194,195]. Polymorphisms of the CCL5 chemokine rs2107538, rs2280788 and rs2280789 had higher frequency in infants with severe RSV-induced bronchiolitis compared to the controls[159]. Infants with RSV-induced bronchiolitis also were found to have higher frequency of CD14 rs5744455 polymorphism[166]. Moreover, the diagnosis of wheeze following RSV-induced LTRI was significantly lower in infants possessing the AA genotype of the rs1800872 polymorphism of IL10 gene[179]. Furthermore, infants with IL13 rs20541 polymorphism had higher frequency of wheeze following RSV-induced bronchiolitis[189].

One of the largest studies conducted on genetic signals underpinning RSV infection was conducted by Janssen *et al*. The genetic association study analysis of 384 SNPs in 220 candidate genes was conducted by Janssen *et a*[208]. Polymorphisms associated with the innate immune genes such as JUN, IFNA5, NOS2A and VDR showed the strongest association with RSV bronchiolitis[208]. Furthermore, genes associated with innate immunity including IFNA5 rs10757212, rs643070 IFNA13, rs2254514 IL15, rs7747909 IL17, rs2274064 NCF2, rs1060826 NOS2, rs1914408 STAT1, rs2407992 TLR8, rs1799724 TNF and rs10735810 VDR[208]. Another study utilised the whole-exome sequencing (WES) approach[374]. WES implements targeted sequencing to capture 95% of the protein-encoding parts of the genome[375]. This potentially aids the discovery of causative genetic signals[375]. Salas *et al* analysed the whole exome of 54 individuals with RSV infection[374]. In comparison to controls, a wide variety of genetic signals were identified[374]. These included polymorphisms of the olfactory gene rs199665292 OR13C5, HLA gene polymorphism rs1047985 and rs935, MUC4 genes rs201623571,

rs529417345 and rs548345415, rs231216 polymorphism of the proline and serine encoding PROSER3 gene, JARID2-encoding gene polymorphism rs72837837 and SNAI3-AS1 gene polymorphism rs3794628[374].

The COVID-19 pandemic led to a great focus of research on host genetic factors associated with SARS-CoV-2 pathogenesis. Several GWAS studies on COVID-19 susceptibility and severity have been conducted. One of the first GWAS studies on COVID-19 host genetic factors was led by Ellinghaus *et al*. Genome-wide significance for SARS-CoV-2 susceptibility and Z (SLC6A20) and rs657152 (ABO) was reported[376]. GWAS by Niemi *et al* found genetic signals associated with COVID-19 susceptibility: rs2271616 (SLC6A20), rs11919389 (PPL24), rs912805253 (ABO), and rs4801778 (PLEKHA4)[377]. Whereas the following genetic signals were associated with COVID-19 severity: rs10490770 (LZTFL1), rs1886814 (FOXP4), rs72711165 (TMEM65), rs10774671 (OAS1-OAS3), rs77534576 (TAC4), rs1819040 (KANSL1), rs74956615 (RAVER1), rs2109069 (DPP9), and rs13050728 (IFNAR2)[377]. Several loci identified by Niemi *et al* have previously been reported to show association with lung or inflammatory or autoimmune disease[377]. Moreover, GWAS by Pairo-Castineira *et al* identified the following genetic signals associated with critical COVID-19 infection: rs73064425 (LZTFL1), rs9380142 (HLA-G), rs143334143 (CCHCR1), rs3131294 (NOTCH4), rs10735079 (OAS1-OAS3), rs2109069 (DDP9), rs74956615 (TYK2), rs2236757 (IFNAR2)[378].

Several case-control candidate studies on COVID-19 susceptibility and severity have been conducted. The influence of the ACE, ACE2 and TMPRSS2 polymorphisms on COVID-19 have been investigated by several studies. Individuals with the I/D genotype were found to have an increased risk of developing symptomatic COVID-19 in a study by Hubacek *et al*[379]. Whereas Verma *et al* reported that the frequency of the D/D genotype and the D allele were significantly higher in COVID-19 patients with severe disease[380]. Similarly Rezaiezadeh *et al* found that severe COVID-19 was associated with the DD genotype of rs4646994 (ACE)[381]. A study by Mohlendick *et al* found that G/G genotype or G allele was associated with a three-fold increased risk in developing a severe COVID-19 disease or a two-fold increased risk of

SARS-CoV-2 infection[382]. Furthermore, association between SARS-CoV-2 infection and the C/C genotype of rs383510 (TMPRSS2) was reported by Schönfelder *et al*[383]. Several risk alleles for susceptibility and severity of COVID-19 infection have been identified. For example, rs12252 polymorphism of the IFITM3 gene showed association with COVID-19 disease severity[384]. Furthermore, the C allele of rs1800795 (IL6) was significantly associated with disease severity in a study by Verma *et al*[385]. A study by Nia *et al* showed that the frequency of G allele of rs909253 (TNF-β) was elevated among COVID-19 patients compared to controls[386]. Whereas the A allele of rs1800629 (TNF-α) was reported to decrease the risk of COVID-19[386]. However, the A allele was reported to increase the risk of COVID-19 in a study by Saleh *et al*[387]. Few studies have reported protective the effect of genetic signals towards SARS-CoV-2 infection. Polymorphism of the APOE gene rs7412, which is a genetic determinant for lipoprotein and lipid levels, was found to confer a protective effect against SARS-CoV-2 infection[388]. Furthermore, the frequency of T allele of rs35705950 (MUC5B) was lower in Dutch COVID-19 patients suggesting protection against the infection[389].

Determination of the candidate causal genes underpinning the genetic signals facilitates analysis of pathways and diseases. Genes and upstream regulators associated with the pathways and diseases can then be targeted to further decipher the mechanistic background behind viral infection. Furthermore, the candidate casual genes can be used as biomarkers and be used for targeted interventions in high-risk individuals. The genes may also serve as druggable targets to facilitate drug development or repurpose drugs to treat viral infection. The lack of efficacy of drugs entering clinical development remains a big problem with over a quarter of drugs failing[390]. This failure in predicting efficacy in humans has prompted developers to use genetic evidence to inform decision making[390]. Several genes associated with disease traits have been identified to be effective druggable targets such as 3-Hydroxy-3-Methylglutaryl-CoA reductase (HMGCR) which are targeted by statins[390]. It is imperative to develop drugs to treat RV, RSV and SARS-CoV-2 infection. Treatment of RV infections remain limited with there currently being no licensed vaccine. The two RSV vaccines that have been approved, GSK's

Arexvy™ and Pfizer's Abrysvo™, are only licensed for use in individuals aged 60 or over[99,100]. Therefore, it is essential to develop an RSV vaccine for infants. Treatments for COVID-19 include vaccine such as BNT162b2 (Comirnaty®; BioNTech, Pfizer)[391], mRNA-1273 (Moderna)[392] and ChAdOx1nCoV-19 (AZD1222; AstraZeneca)[393], and anti-viral drugs such as remdesivir[394]. Patients hospitalised with severe or critical conditions also be treated with immunomodulatory drugs for example dexamethasone[394]. However, not every individual can be given the vaccine and the anti-viral drugs. Furthermore, treatment for long COVID is still needed and it may be useful to develop therapies targeting different coronaviruses. It remains a challenge to understand the molecular basis of genetic associations and so insight into host susceptibility and information for developing new anti-viral treatments can be given by identifying the candidate causal genes and pathways.

3.2 Aims and Objectives

Overall Aim: To complete a signal-to-gene/pathway analysis by interrogating multiple datasets and a drug interactions analysis to identify therapeutic opportunities.

The objectives of the projects are:

- To conduct a literature search to obtain genetic signals thought to influence host-RV/ RSV and SARS/CoV-2 susceptibility and severity.
- To identify candidate causal genes and pathways using signal-to-gene mapping.
- To identify opportunities for new drug development and drug repurposing.

3.3 Methods

Systematic literature review was conducted using different databases. Genetic signals associated with host-RV/RSV/SARS-CoV-2 were identified and scored for genes. Genes scoring 3 or higher were interrogated for pathways, diseases, upstream regulators, and drugs-gene interactions.

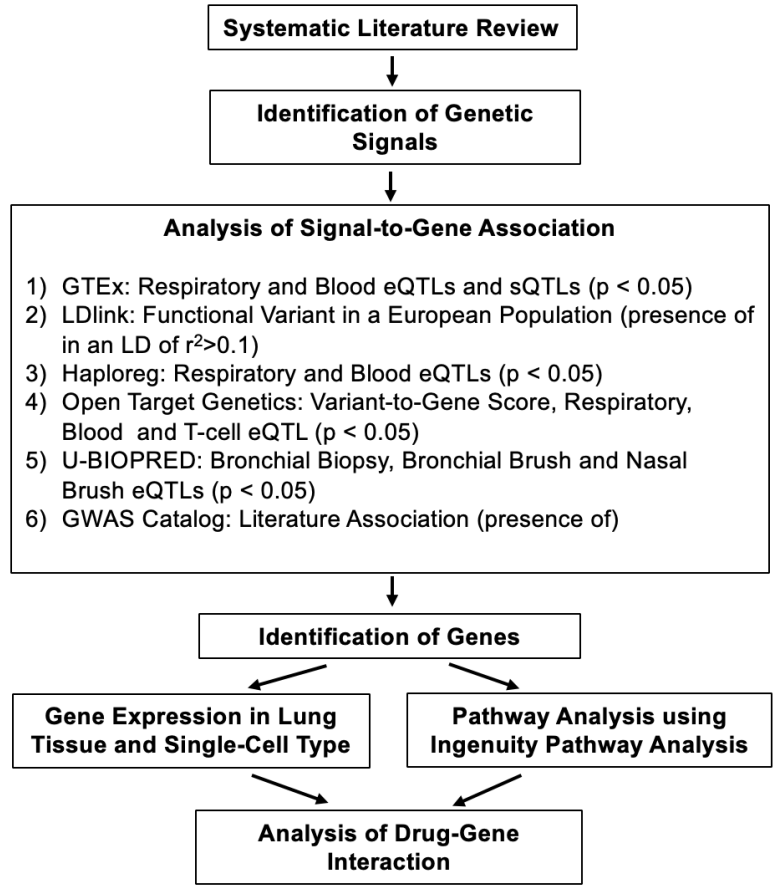
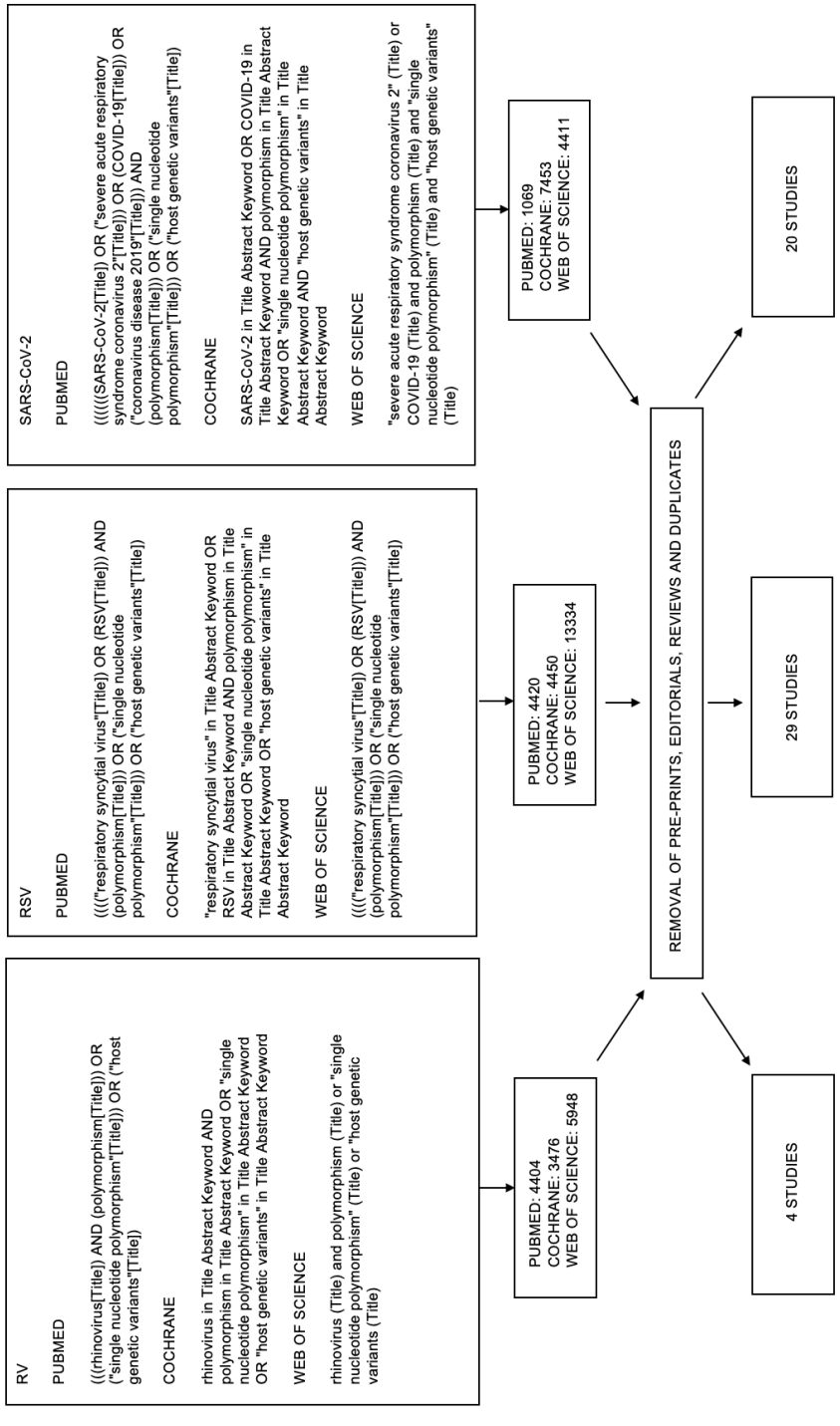


Figure 3-1. Work-flow of the methodology used for signal-to-gene association analysis. Systematic literature review to identify genetic signals. Genes scoring 3 or higher were taken for downstream analysis.

3.3.1 Literature Search of SNPs Associated with Host-Virus Interactions

PubMed, Cochrane Library and Web of Science was used to literature search for SNPs associated with host-RV/RSV/SARS-CoV-2 interactions and obtain genetic signals. For RV and RSV-host interactions, studies were filtered from 01/01/2000-22/05/2023 and 01/01/2019-22/05/2023 for SARS-CoV-2. The search string used is shown below. List of studies were downloaded from PubMed, Cochrane Library and Web of Science and exported onto excel. Papers were screened by title and abstracts. Full text of papers that reported association between SNPs and RV/RSV/SARS-CoV-2 susceptibility or severity were screened. Screening and study selection was single-screen. The systematic literature review has not been registered on PROSPERO registry. Software and AI tools were not used to process the papers. Selected studies were primary research studies with well-defined methodologies including GWAS and case-control studies. The participants involved human subjects with clearly defined sample size and study population. Reviews, duplicates, pre-prints, and editorials were excluded. Pre-prints were excluded as they are not peer-reviewed. Additional literature was searched by checking the references of studies included in the analysis. Due to limited number of data available, studies with small sample size were not excluded. Therefore this could be a limitation as studies with small sample size may not be as representative of the population.

Figure 3-2. Work-flow of the literature search. PubMed, Cochrane Library, and Web of Science was used for the literature search. Literature from 01/01/2000-22/05/2023 for RV and RSV and 01/01/2019-22/05/2023 for SARS-CoV-2 was identified.



3.3.2 Analysis of Signal-to-Gene Association

The following databases were used to score the genetic signals: GTEx V8, LDlink Version, Haploreg Version 4.1, Open Target Genetics (OTG), Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes (U-BIOPRED) and GWAS Catalog (https://www.ebi.ac.uk/gwas/). GTEx, OTG, Haploreg and U-BIOPRED was used to identify expression quantitative trait loci (eQTLs). eQTLs are genetic loci that are associated with expression traits[395]. Mapping of eQTLs identifies chromosomal regions and genetic variants that are associated with phenotypes such as height, eye colour etc[395]. Microarray technologies can simultaneously measure the expression levels of tens of thousands of genes allowing the detection of the genomic regions that regulate these expression levels[395]. The Genotype-Tissue Expression (GTEx) is an atlas of eQTLs spanning 15,201 RNA-sequencing samples from 49 tissues from 838 post-mortem donors[396]. Genetic associations for gene expression and cis and trans splicing are characterised[396]. GTEX helps deepen our understanding on how genetic effects on the transcriptome can facilitate trait associations[396]. Haploreg allows the examination of polymorphisms and small indels and their predicted chromatin state in different cell types[397]. This facilitates analysis of enrichment of enhancers that are cell specific[397]. OTG utilises fine-mapping to systemically prioritise genes to causal variant[398]. OTG is an integrative platform that aggregates functional genomics data such as gene expression, chromatic interaction, protein abundance from a variety of different cell types and tissues[398]. OTG was used to identify eQTLs in lung, whole blood and T-cell, and variant-to-gene scores[398]. U-BIOPRED is a consortium spanning multiple institutions, pharmaceutical companies, and patient organisations[399]. U-BIOPRED was used to identify eQTLs in specialised tissues including bronchial, brushes, bronchial biopsy, and nasal brushes[399].

LDlink is a linkage disequilibrium (LD) analysis tool that allows querying of pairwise LDs between SNPs. LD is a non-random association between alleles at two or more loci in a given population[400]. LD is utilised to map genes associated with quantitative characters[400]. LDlink integrates reference population sets, functional annotations, and interactive output to investigate

functional variants in high LD[401]. LDproxy of the LDlink was used to identify functional variants for the genetic signals[401]. This enables fine mapping of disease susceptibility regions and characterisation of functional variants based in genotype-phenotype associations[401]. Using the LDproxy of LDlink, proxies with the highest R2 value with the minimum of 0.2 were obtained. To note eQTLs, p-value of < 0.05 was used. GWAS Catalogue is a curated collection of published GWAS which assists identification of casual variants, interrogation of disease mechanisms and establishing targets for therapeutics[402]. GWAS Catalog was used to identify published associations[403]. GWAS traits were considered positive if a respiratory or blood/immune trait was associated with a risk allele of a genetic signal at genome-wide significance. A score of 1 was given for each signal-to-gene association identified. Signal-to-gene association obtained from the U-BIOPRED dataset were given a score of 2 as the eQTLs are specific to respiratory related samples. Genes with a score of s ≥ 3 were taken for downstream analysis.

Table 3-1. Methodology for the scoring of the gene from signal. Genes scoring higher than 3 were used in the downstream analysis.

Evidence	Weighting
GTEX Respiratory Relevant Cell/Tissue eQTL (p <0.05)	1
GTEX Respiratory Relevant Cell/Tissue sQTL (p <0.05)	1
GTEX Blood eQTL (p <0.05)	1
GTEX Blood sQTL (p <0.05)	1
Functional Variant in a European Population (presence of in an LD $r^2 > 0.1$)	1
Haploreg Respiratory eQTL (p <0.05)	1
Haploreg Blood eQTL (p <0.05)	1
Open Target Genetics Variant-to-Gene Score (presence of)	1
Open Target Genetics Respiratory eQTL (p <0.05)	1
Open Target Genetics Blood eQTL (p <0.05)	1
Open Target Genetics T-cell eQTL (p <0.05)	1
U-BIOPRED Bronchial Biopsy eQTL (p <0.05)	2
U-BIOPRED Bronchial Brush eQTL (p <0.05)	2
U-BIOPRED Nasal Brush eQTL (p <0.05)	2
Literature Association using GWAS Catalog (presence of)	1
Total Possible Score	**18**

3.3.3 Analysis of PheWAS Using GeneATLAS UK Biobank Data

Phenome-wide association studies (PheWAS) allow identification of association between a phenotype and a specific SNP[404]. PheWAS utilises a broad range of curated human phenotypes (phenome) to explore associations[404]. GeneATLAS database is a platform of genetic associations for 660 binary and 118 non-binary traits of 452,264 participants from UK Biobank of European ancestry[405]. Genome-wide association of 9,113,133 genetic variants can be queried[405]. GeneATLAS was used to explore traits showing association with genetic signals obtained from the literature search[405]. A Bonferroni adjusted p-value of 6.42×10^{-5} was used as there are 778 traits in the database[405]. Traits were selected and grouped into respiratory, blood/immune, allergy, inflammatory, auto-immune and viral. Proxies with the highest R^2 values were used using the LDlink database when the reported variants were unavailable. The following proxies were used: rs35761398 (rs2501432, $R^2 = 1$), rs1047985 (rs9272326, $R^2 = 0.4023$), rs1059047 (rs4253516, $R^2 = 0.9023$), rs14701 (rs13168774, $R^2 = 0.8854$), rs1819040 (rs242561 $R^2 = 0.88$), rs11385942 (rs35508621, $R^2 = 1$), rs9271609 (rs28366298, $R^2 = 0.918$), rs56106917 (rs11610045, $R^2 = 0.996$), rs2532300 (rs242559, $R^2 = 0.82$).

3.3.4 Analysis of Total and Single Cell RNA Expression in Lung Tissue

The Human Protein Atlas (HPA) is a database consisting of protein localisation and expression data in different human tissues and cells[406]. The entire human proteome was mapped using immunohistochemistry and immunocytochemistry, which generated over 10 million images of patterns of protein expression at a single-cell level[406]. This approach coupled with transcriptomics data allows interrogation of global expression profiles[406]. The expression of candidate casual genes in lung tissues were identified using the HPA. The average transcript per million (TPM) values from individual lung samples from three different transcriptomic datasets (HPA, FANTOM5, GTEx) are normalised to generate NX values for total lung expression[406]. Transcripts per million protein coding genes (pTPM) values were used for gene expression

of single cell lung tissue. Read counts from four different lung tissue datasets (the Human Cell Atlas, Single Cell Expression Atlas, the European Genome-phenome Archive, the Gene Expression Omnibus (GEO) were normalised to provide the pTPM values[406]. In the chart showing total expression, genes displaying enriched/enhanced expression blood/immune cells were colour coded red, enrichment/enhancement for lung were colour coded dark blue, enrichment/enhancement for blood/immune and lung were coded pale blue and no enrichment for blood/immune and lung were colour coded black. Log_{10}(pTPM) values were used for the heatmap.

3.3.5 Analysis Using Ingenuity Pathway Analysis (IPA)

The QIAGEN Ingenuity Pathway Analysis facilitates identification of upstream regulatory molecules and their downstream effects via associated mechanisms[407]. IPA aims to elucidate whether biological mechanisms that underlying data are relevant to previously reported causal relationships in literature[407]. List of candidate causal genes for each virus was uploaded onto the database. Canonical pathways were used to identify significant pathways, gene expression patterns, and changes in the activity of the pathways. Significantly associated diseases, upstream regulators and their associated downstream molecules were identified.

3.3.6 Analysis of Drug Gene interaction

The Drug Gene Interaction Database (DGIdb) is a platform that delivers information about drug-gene interactions and druggable gene targets using databases, publications, and web-based sources[408]. DGIdb gives an interaction score which is calculated by taking into account the number of genes that interact with the drug, evidence based on the number of publications and sources supporting the interaction, and the extent to which the drug is known to interact with other genes[408]. Drugs with an interaction score of > 1.0 were noted. Drugs targeting the candidate casual genes which are currently in phase II , phase II and phase IV of clinical trials were noted using clinicaltrial.gov. The ClinicalTrial.gov registry reports information on clinical trials including the purpose of the study, recruitment status, eligibility criteria, location, design of the study, and important details of the protocol[409].

3.4 Results

3.4.1 Analysis of Signal-to-Gene Association Identifies 6, 108 and 119 Candidate Causal Genes Associated with Host-RV/ RSV/ SARS-CoV-2 Interactions Respectively

Through a systematic literature search, manuscripts were collated that identified genetic signals found to be associated with susceptibility to or severity of respiratory viral infections. Three different search engines were used: PubMed, Cochrane Library, and Web of Science. A list of sentinel SNPs was compiled, and signal-to-gene analysis were conducted where each SNP was scored based on previously published methodology[410]. Sentinel SNPs were defined if they were significantly associated with susceptibility or severity to RV/RSV/SARS-CoV-2 disease. 6, 108 and 119 candidate causal genes were translated from 4, 65 and 56 sentinel SNPs for RV, RSV and SARS-CoV-2 respectively. It is possible to have more than one genes translated from a genetic signal. Through functional linkage, SNPs in core regulatory region can alter the binding of transcription factors and simultaneously affect the expression of multiple neighbouring genes and genes further away[411]. The highest score for RV were signals tagged to rs1800890 (IL10; score of 4) and rs1800795 (IL6-AS1). The signals tagged to rs1047985 (HLA-DQA1; score of 15) and rs3131294 (HLA-DQB1; score of 14) scored the highest for RSV and SARS-CoV-2 respectively. There was a degree of overlap between certain genes with 6, 21, and 3 genes showing overlap for RV/RSV, RSV/SARS-CoV-2, and RV/SARS-CoV-2 respectively (Table 3-2).

Table 3-2: Overlap of Candidate Causal Genes between RV, RSV and SARS-CoV-2. Genes that showed overlap between RV/RSV/SARS-CoV-2 are colour-coded orange, RV/RSV are colour-coded green, and RSV/SARS-CoV-2 are colour-coded blue.

RV	RSV	SARS-CoV-2
IL10	AIF1	AIF1
IL6	C4A	C4A
IL6-AS1	C4B	C4B
LGALS9	CCHCR1	CCHCR1
STEAP1B	DDX39B	DDX39B
WSB1	HLA-DQA1	HLA-DQA1
	HLA-DQA2	HLA-DQA2
	HLA-DQB1	HLA-DQB1
	HLA-DQB1-AS1	HLA-DQB1-AS1
	HLA-DQB2	HLA-DQB2
	HLA-DRB1	HLA-DRB1
	HLA-DRB5	HLA-DRB5
	IL10	IL6
	IL6	IL6-AS1
	IL6-AS1	LTA
	LGALS9	MICA
	LTA	MX1
	MICA	SKIV2L
	MX1	STEAP1B
	SKIV2L	TMPRSS2
	STEAP1B	VARS2
	TMPRSS2	
	VARS2	
	WSB1	

The pro-inflammatory cytokine gene IL6, the long non-coding RNA, interleukin 6 antisense RNA 1 (IL6-AS1) and the membrane component gene STEAP family member 1B (STEAP1B) were identified as potentially being relevant to all three viruses (Figure 3-3B). Cytokine genes such as IL6 and IL10 showed overlap between RV/RSV. HLA genes such as HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB1-AS1, HLA-DRB1, HLA-DRB5, and MHC class I polypeptide–related sequence A (MICA) and genes encoding for complement proteins complement component C4A (C4A) and C4B were found to be common between RSV and SARS-CoV-2. MX1 gene, common for both RSV and SARS-CoV-2, encodes for GTP-binding protein which is implicated in inhibition of viral replication[412]. The gene encoding for allograft inflammatory factor 1 (AIF-1), suggested to have a role in activation and proliferation of endothelial cells, was found to be common for RSV and SARS-CoV-2[413]. Other genes showing overlap for RSV and SARS-CoV-2 included CCHCR1 which is located within the psoriasis susceptibility locus psoriasis susceptibility 1 (PSORS1)[414], and DExD-box helicase 39B (DDX39B), which is involved in the exportation, splicing and translation of mRNA[415]. With an important role in the regulation of metastatic diseases in hormone receptor-negative breast cancer, WSB1 was found to overlap between RV and RSV[416]. The serine protease encoding TMPRSS2, which is involved in the cellular entry of SARS-CoV-2, was also identified in RSV susceptibility along with SARS-CoV-2[417]. Another gene identified as being common between RV and RSV was LGALS9, which is important in the regulation of immune response and inflammation[418]. Valyl-TRNA synthetase 2 (VARS2), which encodes for mitochondrial aminoacyl-tRNA synthetases, showed overlap between RSV and SARS-CoV-2[419].

Although genes common for different viruses were identified, many of the genes were found to be exclusively associated with a single virus. This is not surprising, as the structure and nature of these viruses differ. Some of the genes found to be exclusively associated with RSV are implicated in immune response. This includes chemokine CCL5 and chemokine receptor CCR5, cytokines IL4, IL8, IL13, IL15 and IL18, and immunomodulatory mediators IFNG and CD14. Metalloproteinases such as ADAM33, MMP1, and MMP16 were also implicated in RSV susceptibility and severity. ADAM33 plays an

important role in the modelling of the airways[420] whereas MMP16 is thought to be implicated in alveoli development[421]. Human collagenase MMP1 regulates cell behaviour and is involved in wound repair, tissue morphogenesis and developmental processes[422]. Moreover, genes implicated in neurodegenerative disorders were found to be associated with RSV susceptibility and severity. Mutation of the annexin A11 protein (ANXA11) gene, which encodes for Ca^{2+} dependent phospholipid-binding protein, is implicated in the motor neuron disease amyotrophic lateral sclerosis[423,424]. BAG cochaperone 6 (BAG6) detects proteolytic fragments with exposed hydrophobicity and prevents them from aggregating[424]. Furthermore, genes associated with mitochondrial function were identified. For example, mitochondrial catalysis-associated gene cytochrome C oxidase subunit 6B1 (COX6B1), is associated with RSV susceptibility and severity. A subunit of COX, COX6B1, is involved in the electron transport chain[425]. Whereas disruption of the beta-carotene oxygenase 2 (BCO2) gene, which encodes for the BCO2 enzyme responsible for catalysing carotenoids in the inner mitochondrial membrane, is linked to inflammation and metabolic disorders[426].

Genes involved in a wide range of functions were found to be associated with RSV susceptibility and severity. For example, bone morphogenetic protein 4 (BMP4) plays a key role in embryogenesis and developmental processes such as differentiation, cell growth and apoptosis[427]. Cytochrome B-245 alpha chain (CYBA) encodes the P22(phox) protein, which is an essential component of the superoxide-generating nicotinamide adenine dinucleotide phosphate oxidases and involved in the death of microorganisms[428]. As a transcription factor, elongin A (ELOA) promotes RNA polymerase II transcription elongation rate[429]. Moreover FANCD2 And FANCI associated nuclease 1 (FAN1) is involved in the repair of DNA interstrand cross-link[430]. RAD50 along with meiotic recombination 11 (MRE11A) and nibrin (NBN) have an important role in regulating cellular damage response following DNA double-strand breaks[431]. Another gene associated with RSV susceptibility, ring finger protein 166 (RNF166), has various functions including recruitment of adaptor proteins and ubiquitin-dependent degradation[432,433] whereas septin 8 (SEPTIN8) interacts with mitogen-activated protein kinase 5 to mediate vesicle trafficking[434].

Members of the solute carrier family, SLC22A4 and SLC22A5 were also associated with RSV susceptibility and severity. Another gene, SMG7, is important for dephosphorylating the regulator of nonsense transcript 1 (UPF1) to facilitate nonsense-mediated mRNA decay[435]. Moreover SRA1 modulates the transcriptional activity of steroid receptor[436].

Genes encoding acyl-CoA synthetase family members Acyl-CoA synthetase family member 3 (ACSF3) and ACSL6 were found to be exclusive to SARS-CoV-2. advanced glycosylation end-product specific receptor (AGER) gene, which encodes for the receptor for advanced glycation end-products (RAGE), was unique to SARS-CoV-2[437]. Expressed predominantly in alveolar epithelial cells and macrophages, RAGE has been shown to be implicated in COVID-19-induced lung inflammation[437]. Members of the apolipoprotein APOE and apolipoprotein M (APOM) were found to be unique to SARS-CoV-2. APOE and APPB3 are associated with Alzheimer's disease[438,439]. Phospholipid transporter-encoding gene ATPase phospholipid transporting 11A (ATP11A) was also found to be unique to SARS-CoV-2. Branched-chain amino acid aminotransferases are encoded by the branched-chain amino acid transaminase 2 (BCAT2) gene[440]. BCL11A is essential for hematopoietic stem cell function and regulates the transition from foetal to adult erythropoiesis [441].

Genes encoding several adhesion molecules have been associated with SARS-CoV-2 susceptibility and severity. Adhesion molecules such as Cadherin15 (CDH15) are important for maintaining the structure, function, and plasticity of synapses[442]. Intercellular adhesion molecule 5 (ICAM5) also has function in synaptic development with roles spanning morphogenesis of dendritic spine and higher-order cognitive and motor functions[443]. ICAM1 has numerous essential roles including regulation of leukocyte recruitment to sites of inflammation and epithelial injury-resolution responses[444]. Dimethylarginine dimethylaminohydrolase (DDAH2) regulates RIG-I-like receptor-mediated anti-viral response[445]. Euchromatic histone-lysine N-methyltransferase 2 (EHMT2) is implicated in the methylation of non-histone proteins[446]. The psoriasis susceptibility genes, PSORS1C1 and PSORS1C2, was also associated with SARS-CoV-2 susceptibility and severity. Ras interacting protein 1 (RASIP1) mediates Rho GTPase signalling to regulate stability and

growth of blood vessels[447]. Ribonucleoprotein PTB-binding 1 (RAVER1) is a multidomain protein that forms complexes with microfilament-associated proteins[448]. Members of the solute carrier family (SLC22A31, SLC6A20) and tripartite motif (TRIM) family (TRIM26, TRIM31, TRIM46) were found to be associated with SARS-CoV-2 susceptibility and severity. Transcription factor 19 (TCF19) is an epigenetic reader protein that moderates glucose homeostasis through histone post-translational modifications[449]. A glycoprotein of the extracellular matrix (ECM), tenascin X (TNXB), interacts with various ECM proteins to accelerate fibrillogenesis of collagen[450]. Wnt family member 3 (WNT3) is an essential modulator of embryonic stem cells and embryogenesis[451]. A crucial regulator of genomic imprinting, zinc finger protein 57 homolog (ZFP57), dictates expression of imprinted genes by regulating differential DNA methylation[452]. Another zinc finger, zinc finger GATA like protein 1 (ZGLP1), confers oogenic fate[453].

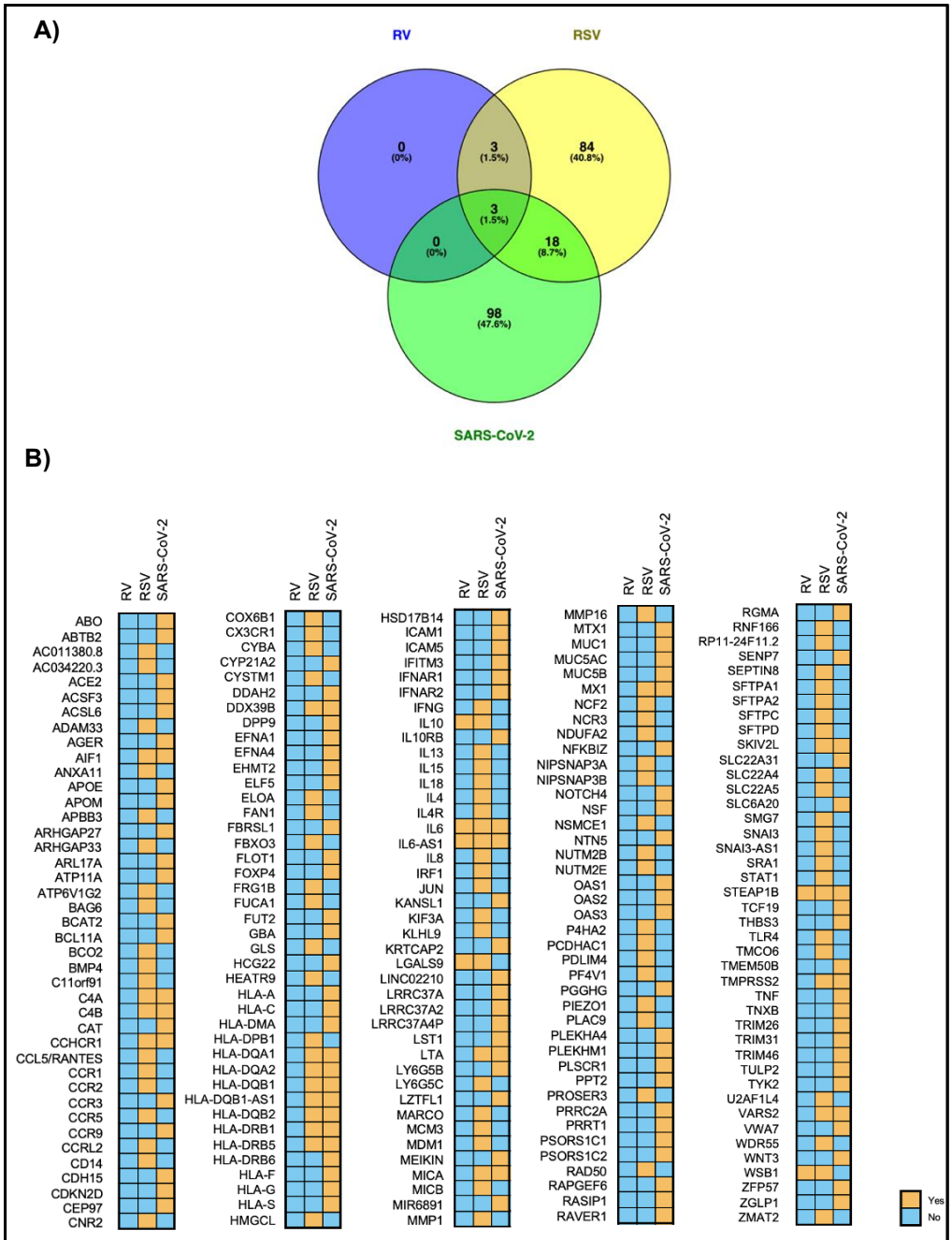


Figure 3-3. A) Venn diagram showing gene overlap identified for Host-RV/ RSV/ SARS-CoV-2 Interactions. B) Schematic to show the gene overlap scoring 3 or higher between the 3 viruses. Orange indicates that the gene is implicated as a candidate causal gene related to virus susceptibility and blue are genes not identified in the signal-to-gene association with that virus.

3.4.2 Cell Type Specificity Was Observed Through Analysis of Single Cell RNA Expression in Lung Tissue

Human Protein Atlas was utilized to explore the expression profile of candidate causal genes in cells of the respiratory system and blood/immune cells. AIF1, and IL10 were particularly enhanced/enriched in blood components (Figure 3-4). Enrichment/enhancement of the following genes were identified in both blood and lung tissues: HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DRB5. C4A, C4B, CCHCR1, DDX39B, HLA-DQB2, LGALS9, LTA, MICA, SKIV2L, STEAP1B, TMPRSS2, VARS2 and WSB1 had no enrichment/enhancement in blood or lung tissues. High expression of IL6 were found in endothelial cells out of the three overlapping genes between the three viruses (Figure 3-4). In comparison to the other cell types, endothelial cells and B-cells expressed greater expression of STEAP1B (Figure 3-4). Single-cell data for IL6-AS1 was unavailable. However, expression for IL6-AS1 in human bronchial fibroblasts was confirmed recently[454]. The expression of HLA genes along with AIF1 were particularly high in macrophages (Figure 3-4).

In comparison to other genes, the expression of C4A, C4B and VARS2 were lower in different lung single cell types (Figure 3-4). The expression of CCHCR1 were higher in macrophages, endothelial cells and ciliated cells compared to the other lung single cell types (Figure 3-4). Similar to CCHCR1, the expression of DDX39B was higher in macrophages and endothelial cells along with T-cells (Figure 3-4). The HLA genes and AIF1 showed high expression in type 2 alveolar cells and macrophages (Figure 3-4). Furthermore, endothelial cells, T-cells and B-cells also showed higher expression of HLA genes (Figure 3-4). The cell type with the highest expression of IL10, galectin-9 (LGALS9) and MX1 were macrophages (Figure 3-4). The expression of LTA was higher in T-cells compared to other cell types. Whereas the expression of helicase SKI2W (SKIV2L) and TMPRSS2 were the highest in macrophages and alveolar cells type 2 respectively (Figure 3-4). The expression of WSB1 was high in a variety of different lung tissue type, such as alveolar cells type 2 and 1, granulocytes, club cells and smooth muscle cells, compared to other genes (Figure 3-4).

IL-10 was enriched/enhanced in blood components for overlapping RV-implicated genes (Figure 3-5). The expression of IL-6 was particularly high in endothelial cells whereas expression of IL-10 and LGALS9 was higher in macrophages for RV (Figure 3-5). For non-overlapping RSV-implicated genes, the surfactant protein SFTPA1, SFTPA2 and SFTPD were particularly enriched/enhanced in lung tissues. Furthermore, the surfactant proteins showed very high expression in alveolar cell type 2 and 1 cells for RSV (Figure 3-6). Genes involved in the immune response, CCL5, Interferon Regulatory Factor 1 (IRF1), the chemokine receptors CCR1, CCR2 and CCR5, and IL8 were enriched/enhanced in blood tissues for RSV (Figure 3-6). With roles in superoxide production, CYBA and NCF2 were also found to be enriched/enhanced in blood tissues[428,455]. RNF166, which is implicated in polyubiquitination of proteins, and SLC22A4, which facilitates elimination of drugs and toxins, exhibited enrichment/enhancement in blood tissues[433]. The macrophage receptor gene MARCO showed enrichment/enhancement in both blood and lung tissues.

Higher expression of CYBA, MARCO, NCF2 genes were identified in macrophages for RSV (Figure 3-6). Genes encoding IL4R, mini-chromosome maintenance protein (MCM3) and placenta associated protein (PLAC9) showed higher expression in endothelial cells compared to the other cell types. The expression of PLAC9 was also high in fibroblasts (Figure 3-6). The immune response genes, CCL5, IFNG and IFR1, showed high expression in T-cells (Figure 3-6). In comparison to other genes, CYBA, cysteine rich transmembrane module containing 1 (CYSTM1), IL18, JUN had a higher expression in granulocytes (Figure 3-6). The expression of IL8, JUN, MMP1 and surfactant proteins were high in club cells (Figure 3-6). COX6B1, CYBA, JUN also showed high expression in B-cells (Figure 3-6). Furthermore, the expression of COX6B1 and CYBA were high in ciliated cells along with JUN. Smooth muscles showed high expression of JUN and PLAC9 (Figure 3-6).

Examining the non-overlapping genes SARS-CoV-2 data shows that AGER, ICAM1 and SLC22A31 is enriched/enhanced in lung tissues (Figure 3-7). Whereas BCL11A, leukocyte-specific transcript 1 protein (LST1), NF-kappa-B inhibitor zeta (NFKBIZ) and TNF showed enrichment/enhancement in blood tissues. The expression of AGER, APOM, ICAM1, IFITM3, LST1, MUC1, NFKBIZ and SLC2AA31 were high in alveolar type 2 cells (Figure 3-7). In comparison to other cell types, the expression of APOE was very high in macrophages. Furthermore, APOM, catalase (CAT), LST1, OAS1 and PLSCR1 also showed high expression in macrophages (Figure 3-7). Expression of AGER and MUC1 exhibited high expression in alveolar type 1 cells whereas IFITM3, NFKBIZ, and NOTCH4 showed high expression in endothelial cells (Figure 3-7). Fibroblasts showed high expression of IFITM3, NFKBIZ and TNXB. NFKBIZ along with APOE and TNF also displayed high expression in T-cells. Moreover, the expression of DDAH2, flotillin-1 (FLOT1) and NFKBIZ were high in granulocytes. Club cells displayed higher expression of IFITM3, MUC1, MUC5AC, MUC5B and NFKBIZ. The expression of BCL11A and cyclin-dependent kinase 4 inhibitor D (CDKN2D) were high in B-cells. Ciliated cells showed high expression of APOM, CAT, IFITM3, LZTFL1, MUC1 and protein-glucosylgalactosylhydroxylysine glucosidase (PGGHG) (Figure 3-7). The levels of APOE, IFITM3, and NFKBIZ were high in smooth muscle cells. These data support the conclusion that the airway structural cells and the subsequent inflammatory cell response may be regulated by the identified susceptibility genes in response to viral infection.

Figure 3-4. Total lung expression and gene expression in single cells of candidate causal genes identified for multiple viruses. Colours of the bars correspond to enrichment of lung, blood/lung and neither. Expression data was cited using Human Protein Atlas proteinatlas.org.

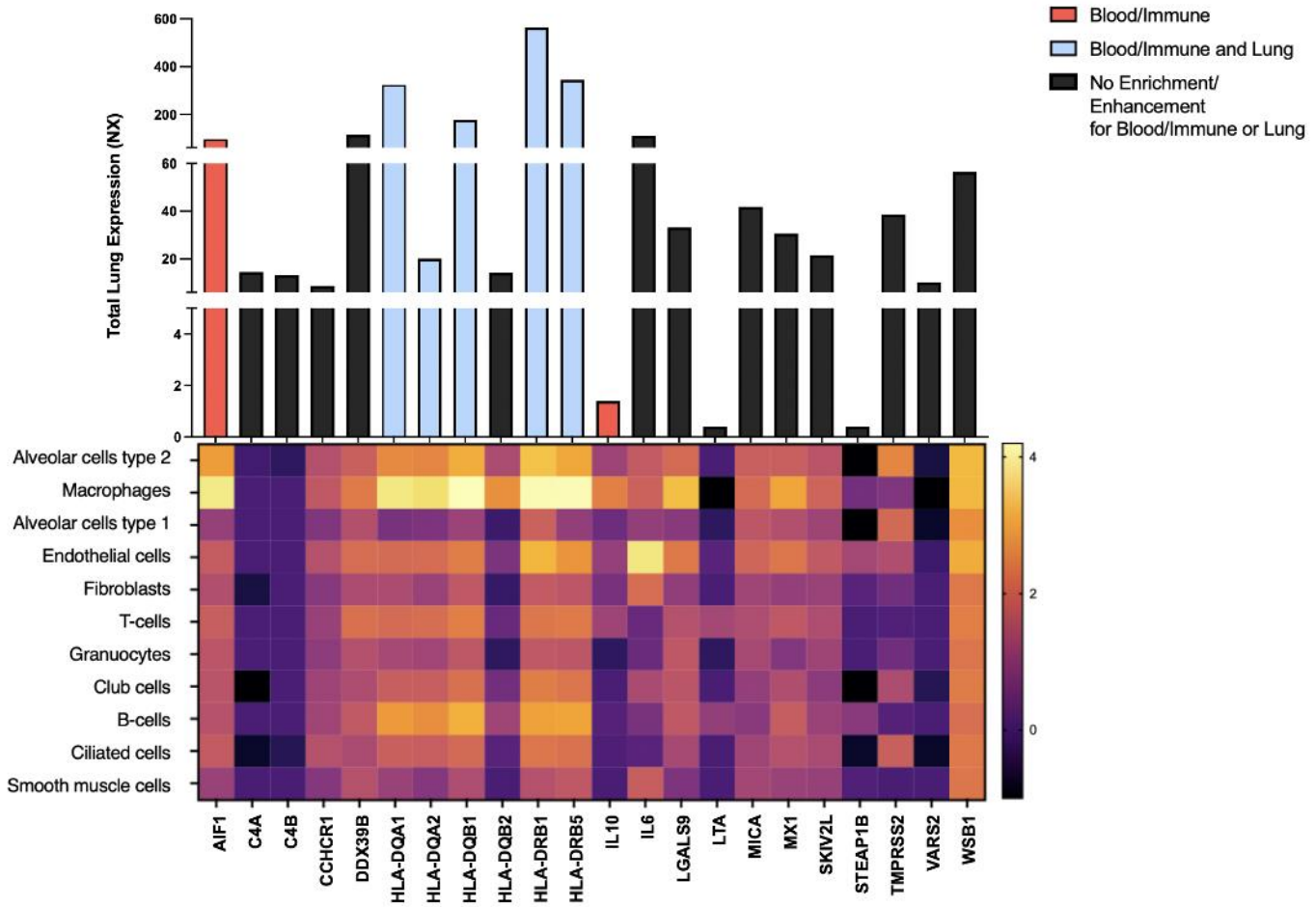


Figure 3-5. Total lung expression and gene expression in single cells of candidate causal genes identified for RV. Colours of the bars correspond to enrichment of lung, blood/lung and neither. Expression data was cited using Human Protein Atlas proteinatlas.org.

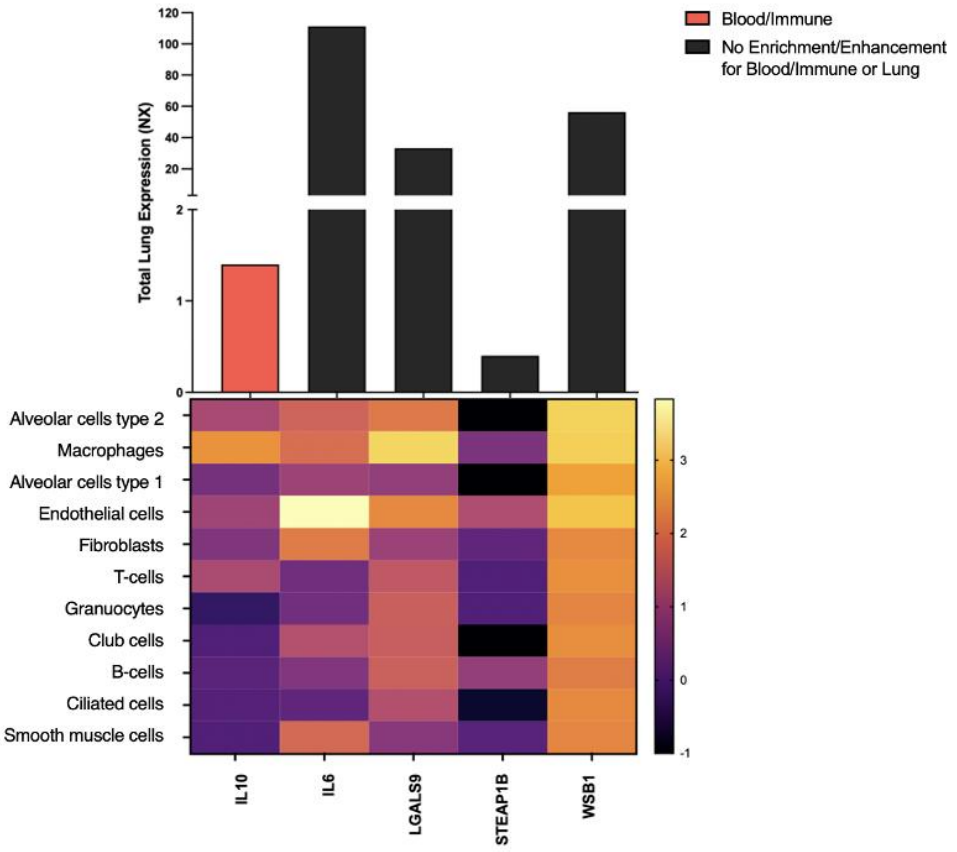


Figure 3-6. Total lung expression and gene expression in single cells of candidate causal genes identified for RSV. Colours of the bars correspond to enrichment of lung, blood/lung and neither. Expression data was cited using Human Protein Atlas proteinatlas.org.

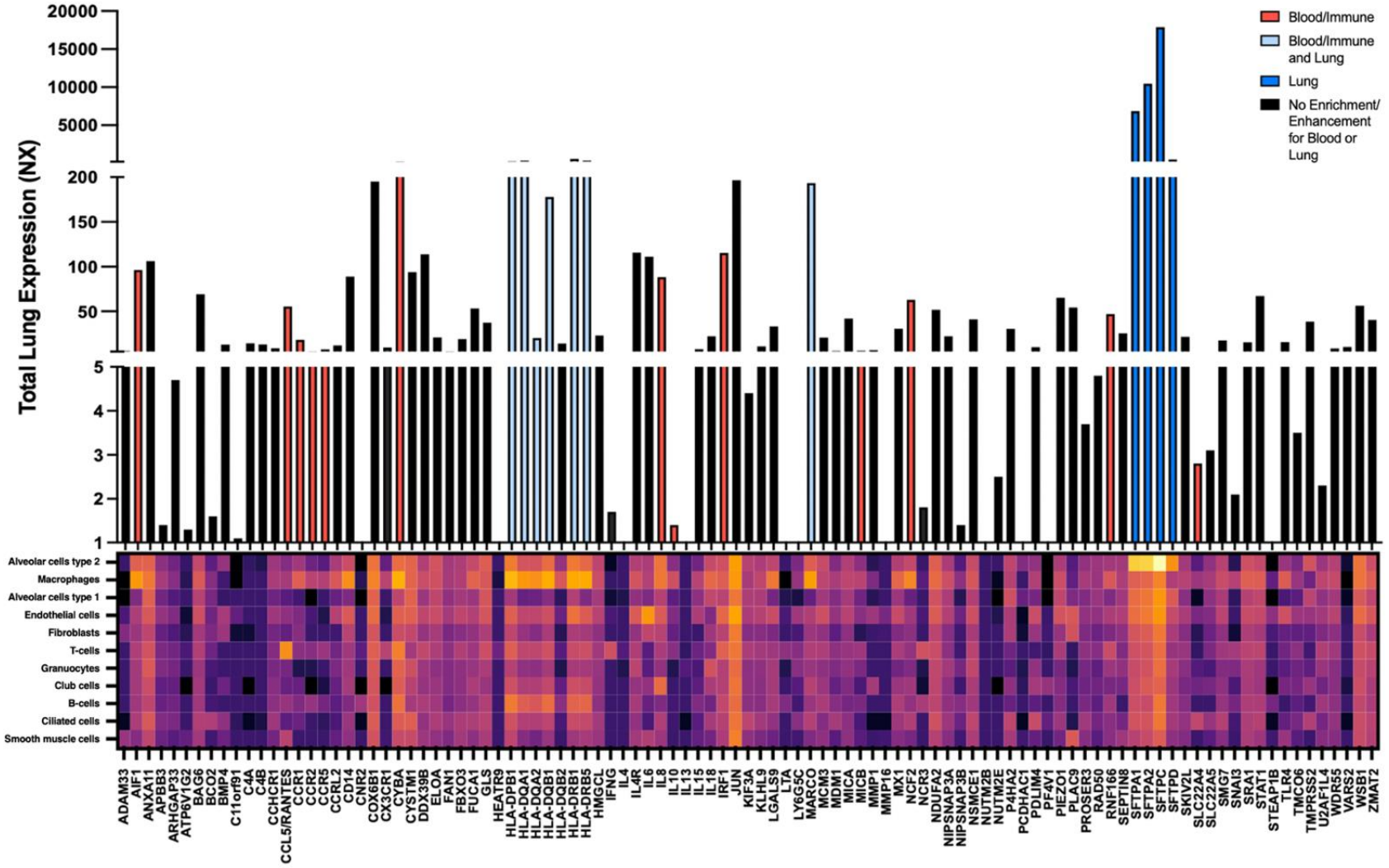
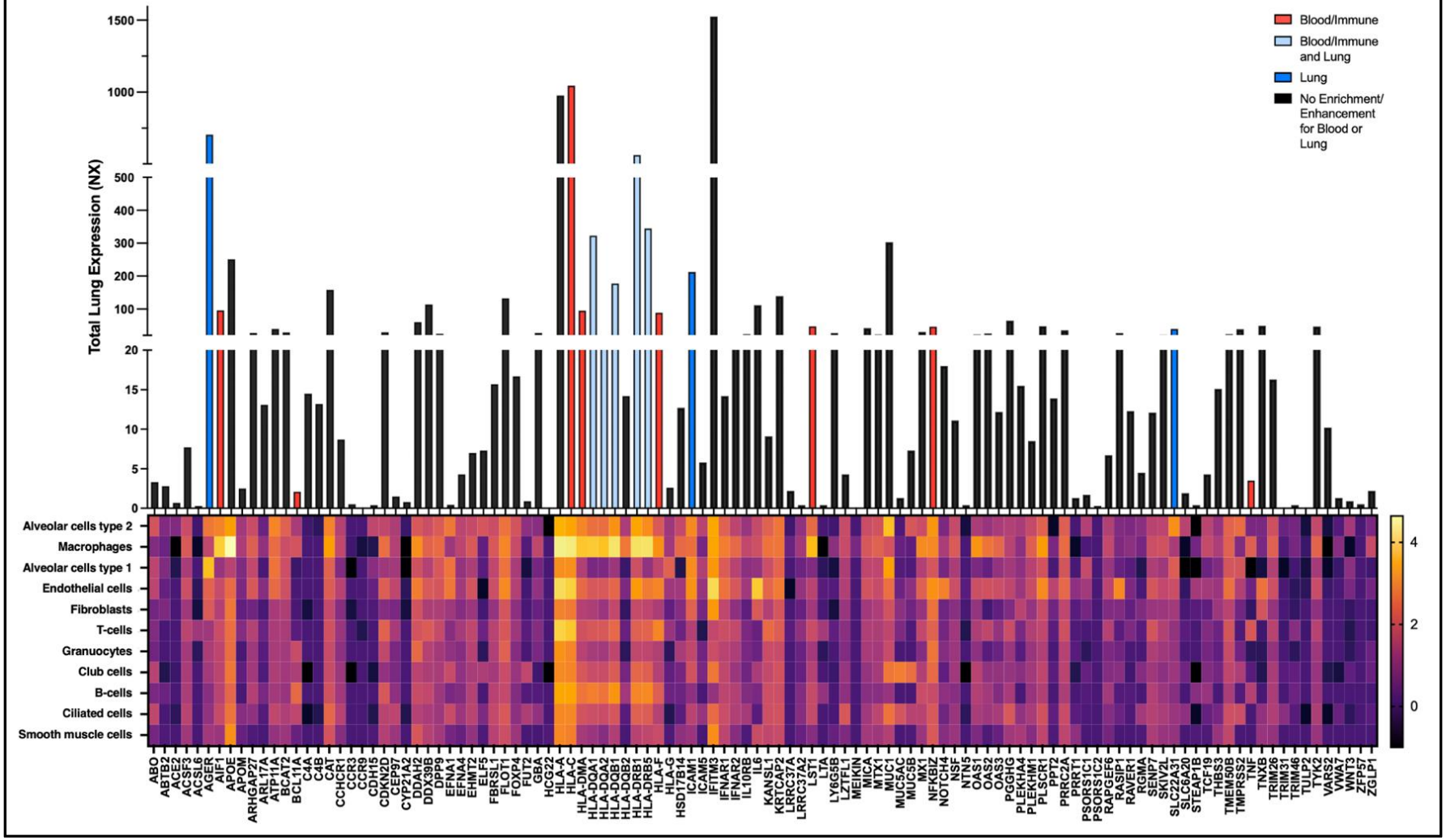


Figure 3-7. Total lung expression and gene expression in single cells of candidate causal genes identified for SARS-CoV-2. Colours of the bars correspond to enrichment of lung, blood/lung and neither. Expression data was cited using Human Protein Atlas proteinatlas.org.



3.4.3 Signals Show Association with Respiratory, Blood Cell Counts, Allergy, Inflammatory and Auto-immune Disorders

PheWAS for 778 traits in the UK Biobank resource was interrogated to gain additional insight into the 14 signals underpinning the 24 candidate causal genes implicated in more than one host-virus interaction (Figure 3-8). The following signals showed association with the respiratory trait asthma, bronchiectasis, nasal polyp, and chronic lower respiratory diseases: rs143334143:CCHCR1, rs1800629:TNF-a, rs1800795:IL6, rs1800890:IL10 rs9271609:HLA-DRB1, rs3131294:NOTCH4, rs909253 TNFB and rs1047985:HLA-DQA1 (Figure 3-8). Allergic traits showed association with the signals rs143334143:CCHCR1 and rs9271609:HLA-DRB1. Allergic traits included allergy/hypersensitive/anaphylaxis, eczema/dermatitis, and hayfever/allergic rhinitis. Psoriasis showed association with the signals rs9271609:HLA-DRB1, rs3131294:NOTCH4, rs909253 TNFB, rs1047985:HLA-DQA1. Found within the psoriasis susceptibility locus PSORS1, CCHCR1 is postulated as a candidate gene responsible for the risk effect of psoriasis[414]. Furthermore, rs9271609:HLA-DRB1 and rs3131294:NOTCH4 showed association with inflammatory conditions such as ulcerative colitis.

Auto-immune conditions also showed associations. Hypothyroidism/ myxoedema showed association with rs143334143:CCHCR1, rs1800629:TNF-a, rs3131294:NOTCH4, rs909253:TNFB and rs1047985:HLA-DQA1. In particular, rs3131294:NOTCH4 showed associated with various auto-immune conditions such as diabetes, hypothyroidism, multiple sclerosis, and rheumatoid arthritis (Figure 3-8). The signals rs1800629:TNF-a, rs9271609:HLA-DRB1, rs3131294:NOTCH4 and rs909253:TNFB showed association with diabetes. Moreover, rs9271609:HLA-DRB1 and rs3131294:NOTCH4 were associated with multiple sclerosis which was also identified in the top 10 pathways for both RSV and SARS-CoV-2. The signal tagged to the three overlapping genes at rs1800795, showed association with the respiratory trait asthma and blood/immune traits such as eosinophil, lymphocyte, and neutrophil count.

For genetic signals unique to viruses, rs1800890:IL10 associated with RV susceptibility showed association with J33 Nasal Polyp and immune traits such as eosinophil percentage and neutrophil percentage (Figure 3-9). Genetic signals associated with RSV susceptibility that have been implicate in the IL4/IL13 axis showed association with asthmatic traits (Figure 3-10). This included rs1800925:IL13, rs1805010:IL4R, rs20541:IL13, rs2070874:IL4, rs2227284:IL4, rs2243250:IL4, rs2243270:IL4, rs734244:IL4 (Figure 3-10). The genetic signal postulated to encode GMSCF (rs7727544) also showed association with asthma (Figure 3-10). Furthermore the following asthma associated- genetic signals also showed association with chronic respiratory disease: rs1800925:IL13, rs1805010:IL4R and rs20541:IL13 (Figure 3-10). Various genetic signals showed association with eosinophil count including rs14701:SLC22A5, rs1799724 TNF, rs1800925:IL13, rs1801275:IL4R, rs1805010:IL4R, rs20541:IL13, rs2070874:IL4, rs2227284:IL4, rs2228570:VDR, rs2243250:IL4, rs2243270:IL4, rs274549:SLC22A5, rs35761398:CNR2, and rs7727544: GMSCF (Figure 10). Association between the allergic trait eczema and dermatitis and the following genetic signals were observed: rs1800925:IL13, rs20541:IL13, rs2070874:IL4, rs2227284:IL4, rs2243250:IL4, rs2243270:IL4 and rs734244:IL4 (Figure 3-10). The genetics signals rs1800925:IL13 and rs20541:IL13 showed association with psoriasis. rs1047985: HLA-DQA1 and rs3732378: CX3CR1 was associated with the auto-immune conditions disorders of the thyroid gland and rheumatoid arthritis (Figure 3-10).

Genetic associations unique to SARS-CoV-2 susceptibility and severity showed association with respiratory traits. Asthma was associated with rs9411378:ABO whereas allergy, hypersensitivity and anaphylaxis was associated with rs73510898:ZGLP1 (Figure 3-11). Associations between inflammatory conditions and several signals were identified. rs9380142:HLA-G showed association with inflammatory bowel disease, ulcerative colitis and psoriasis (Figure 3-11). Furthermore, rs73510898:ZGLP1 was associated with ulcerative colitis and psoriasis and rs34536443:TYK2 was associated with psoriasis (Figure 3-11). Overall, these data suggest that these signals are predominantly implicated in inflammation. Furthermore, a diverse range of

traits were shown to be associated with overlapping signals. This may facilitate understanding of the mechanisms behind pathways and diseases.

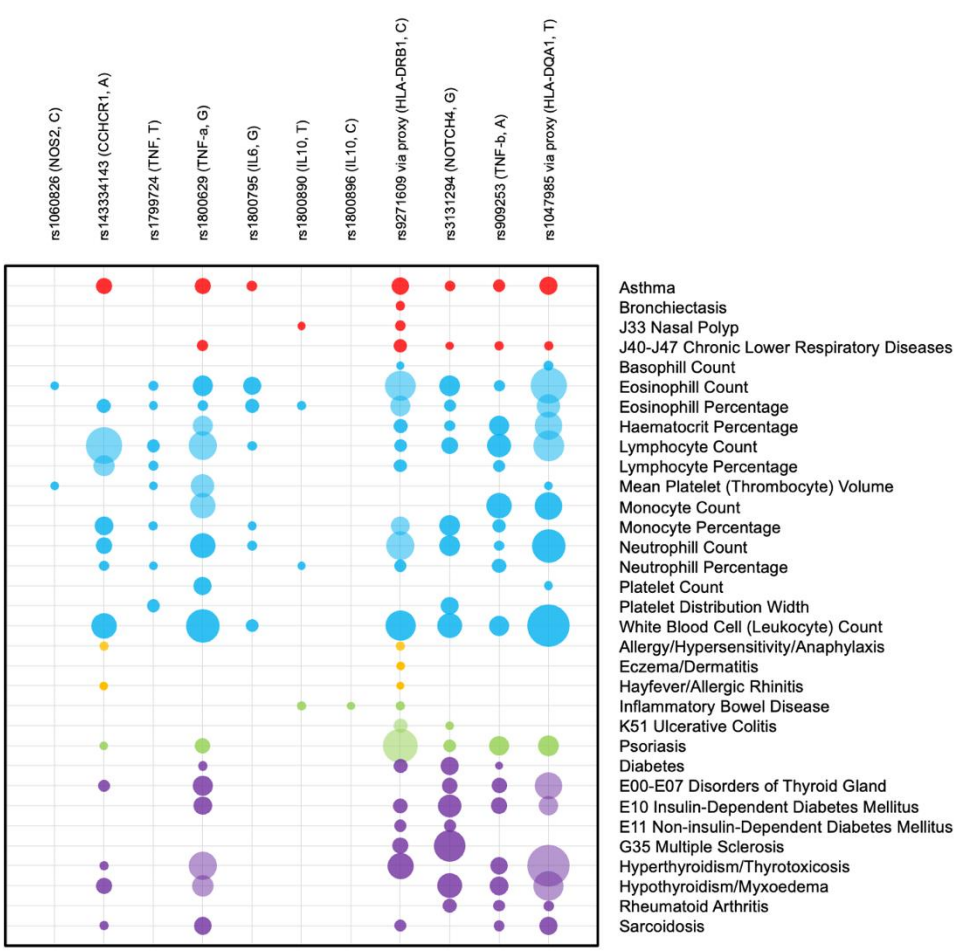


Figure 3-8. PheWAS of genetic signals associated with candidate causal genes that showed overlap between multiple viruses. The p-value size corresponds to the size of the bubbles. Lower p-value corresponds to bigger bubbles. The bubbles are colour-coded based on traits with red:respiratory, blue:blood/immune, yellow:allergy, green:inflammatory, purple:auto-immune.

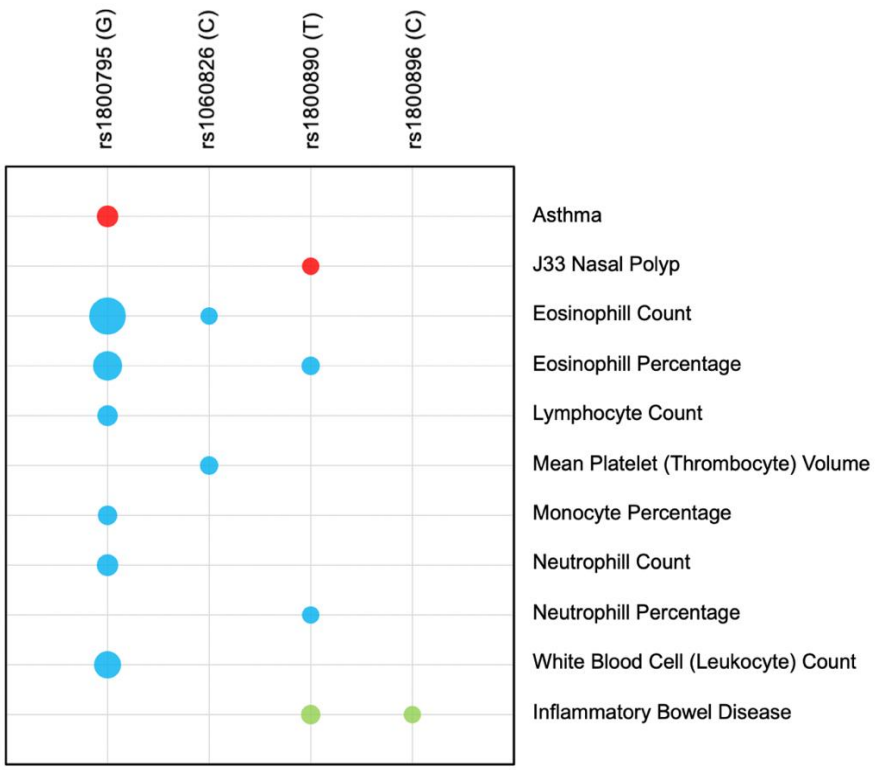


Figure 3-9. PheWAS of genetic signals associated with candidate causal genes for RV. The p-value size corresponds to the size of the bubbles. Lower p-value corresponds to bigger bubbles. The bubbles are colour-coded based on traits with red:respiratory, blue:blood/immune, green:inflammatory.

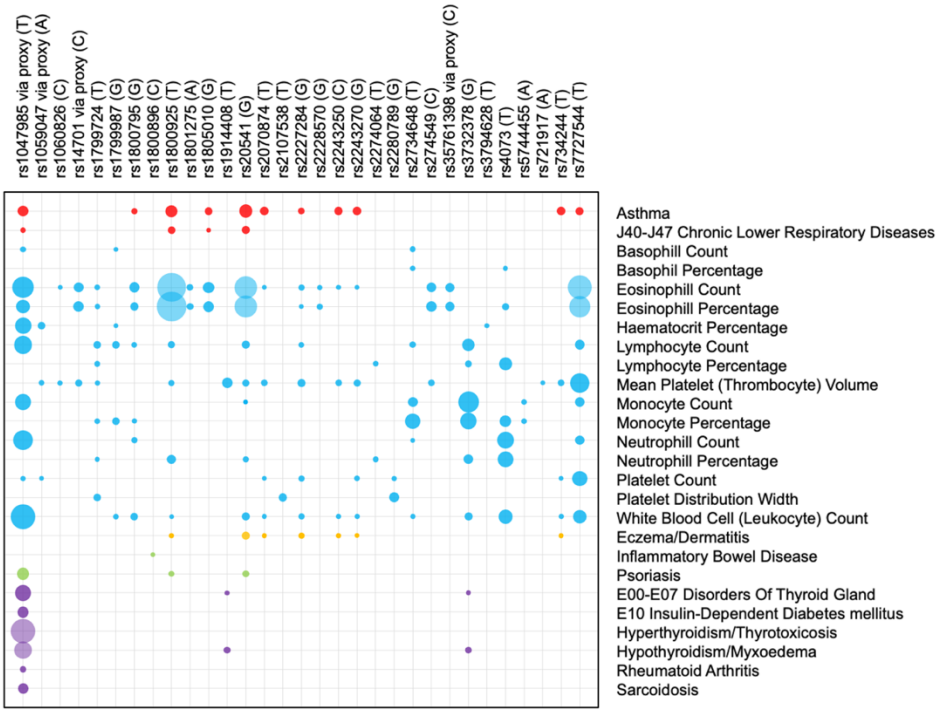


Figure 3-10. PheWAS of genetic signals associated with candidate causal genes for RSV. The p-value size corresponds to the size of the bubbles. Lower p-value corresponds to bigger bubbles. The bubbles are colour-coded based on traits with red:respiratory, blue:blood/immune, yellow:allergy, green:inflammatory, purple:auto-immune.

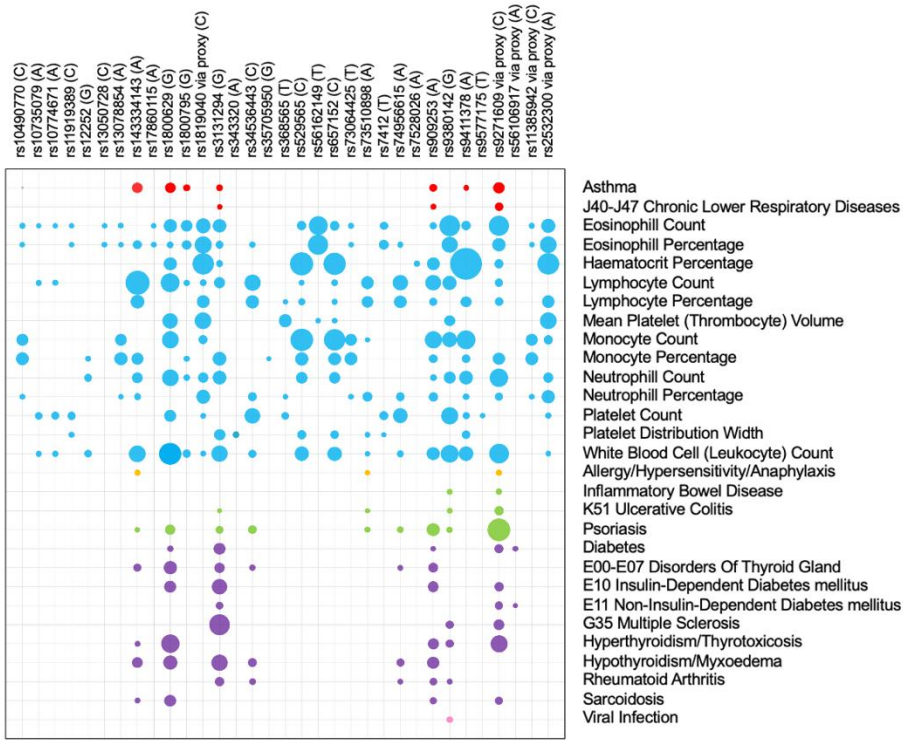


Figure 3-11. PheWAS of genetic signals associated with candidate causal genes for SARS-CoV-2. The p-value size corresponds to the size of the bubbles. Lower p-value corresponds to bigger bubbles. The bubbles are colour-coded based on traits with red:respiratory, blue:blood/immune, yellow:allergy, green:inflammatory, purple:auto-immune, pink:viral.

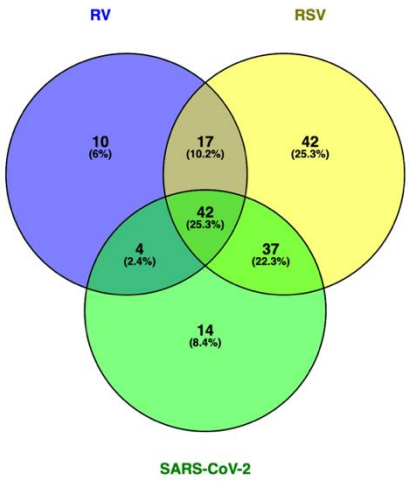
3.4.4 IPA Analysis Identifies Pathways Associated with Host-Virus Interactions

Ingenuity Pathway Analysis (IPA) was used to interrogate the candidate causal genes implicated in the signal to gene mapping. From the candidate causal genes, 73, 138 and 97 pathways potentially mediating susceptibility to infection of RV, RSV and SARS-CoV-2 respectively were identified. Interestingly a significant overlap (25.3%) was observed between all three viruses (Figure 3-12A). This suggests that the immunomodulatory pathway may be particularly important for susceptibility to or severity of RV, RSV and SARS-CoV-2 infection. Common pathways between the three viruses were observed. 59, 79 and 46 pathways overlapped for RV/RSV, RSV-SARS-CoV-2, and RV/SARS-CoV-2 respectively. For all three viruses, 42 pathways were found to be common with pathways diverse as tumor microenvironment pathway and Th1 and Th2 activation pathway being identified. Th1 and Th2 activation pathways and the Th1 pathway showed overlap between all three viruses (Figure 3-12B). Of particular note, the neuroinflammation signalling pathway was identified to be common for both RSV and SARS-CoV-2 (Figure 3-12B). Other overlapping pathways in the top 10 included crosstalk between dendritic cells and natural killer cells, macrophage activation pathway, multiple sclerosis signaling pathway, and Th2 pathway (RSV/SARS-CoV-2) (Figure 3-12B). These findings suggest that a significant immunomodulatory mechanism could be employed by RV, RSV and SARS-CoV-2 upon infection which aligns well with the PheWAS findings.

Pathways implicating the airways were also identified for RV, RSV and SARS-CoV-2 such as airway inflammation in asthma for RV and RSV, and airway pathology in chronic obstructive pulmonary disease for RV, RSV and SARS-CoV-2. Coronavirus pathogenesis pathway was also common between the three viruses. Moreover, 42 pathways were found to be unique for RSV whereas 10 and 14 pathways were unique for RV and SARS-CoV-2 respectively. One of the pathways unique to RV was the Th17 activation pathway. Pathways unique to RSV included the RAR activation and neutrophil extracellular trap signaling pathway. For SARS-CoV-2, pathways found to be

unique included coronavirus replication pathway. Novel pathways implicating autoimmune disorders such systemic lupus erythematosus and rheumatoid arthritis were observed between the three viruses. These include systemic lupus erythematosus in T cell signaling pathway and systemic lupus erythematosus in B cell signaling pathway. For genes identified for RV, Th1 was the top pathway (Figure 3-13). For RSV, the most significant pathway was the macrophage activation pathway (Figure 3-14). The top pathway for SARS-CoV-2 was IL10 signalling (Figure 3-15).

A)



B)

RV	RSV	SARS-CoV-2
Th1 Pathway	Macrophage Classical Activation Signaling Pathway	IL-10 Signaling
Th1 and Th2 Activation Pathway	Th1 and Th2 Activation Pathway	Th1 Pathway
Tumor Microenvironment Pathway	Th1 Pathway	Antigen Presentation Pathway
Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	Pathogen Induced Cytokine Storm Signaling Pathway	Th1 and Th2 Activation Pathway
T Cell Exhaustion Signaling Pathway	Multiple Sclerosis Signaling Pathway	PD-1, PD-L1 cancer immunotherapy pathway
Role of Cytokines in Mediating Communication between Immune Cells	IL-10 Signaling	Multiple Sclerosis Signaling Pathway
Activation of IRF by Cytosolic Pattern Recognition Receptors	Macrophage Alternative Activation Signaling Pathway	Th2 Pathway
TREM1 Signaling	Neuroinflammation Signaling Pathway	Neuroinflammation Signaling Pathway
Role of JAK family kinases in IL-6-type Cytokine Signaling	Th2 Pathway	Macrophage Classical Activation Signaling Pathway
Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza	Crosstalk between Dendritic Cells and Natural Killer Cells	Crosstalk between Dendritic Cells and Natural Killer Cells

Figure 3-12. A) Venn diagram showing pathway overlap for genes implicated in Host-RV/ RSV/ SARS-CoV-2 susceptibility. B) Table showing the top 10 pathways for each RV, RSV and SARS-CoV-2. Pathway is common between all virus is colour coded blue. Pathway common between RSV and SARS-CoV-2 is highlighted orange.

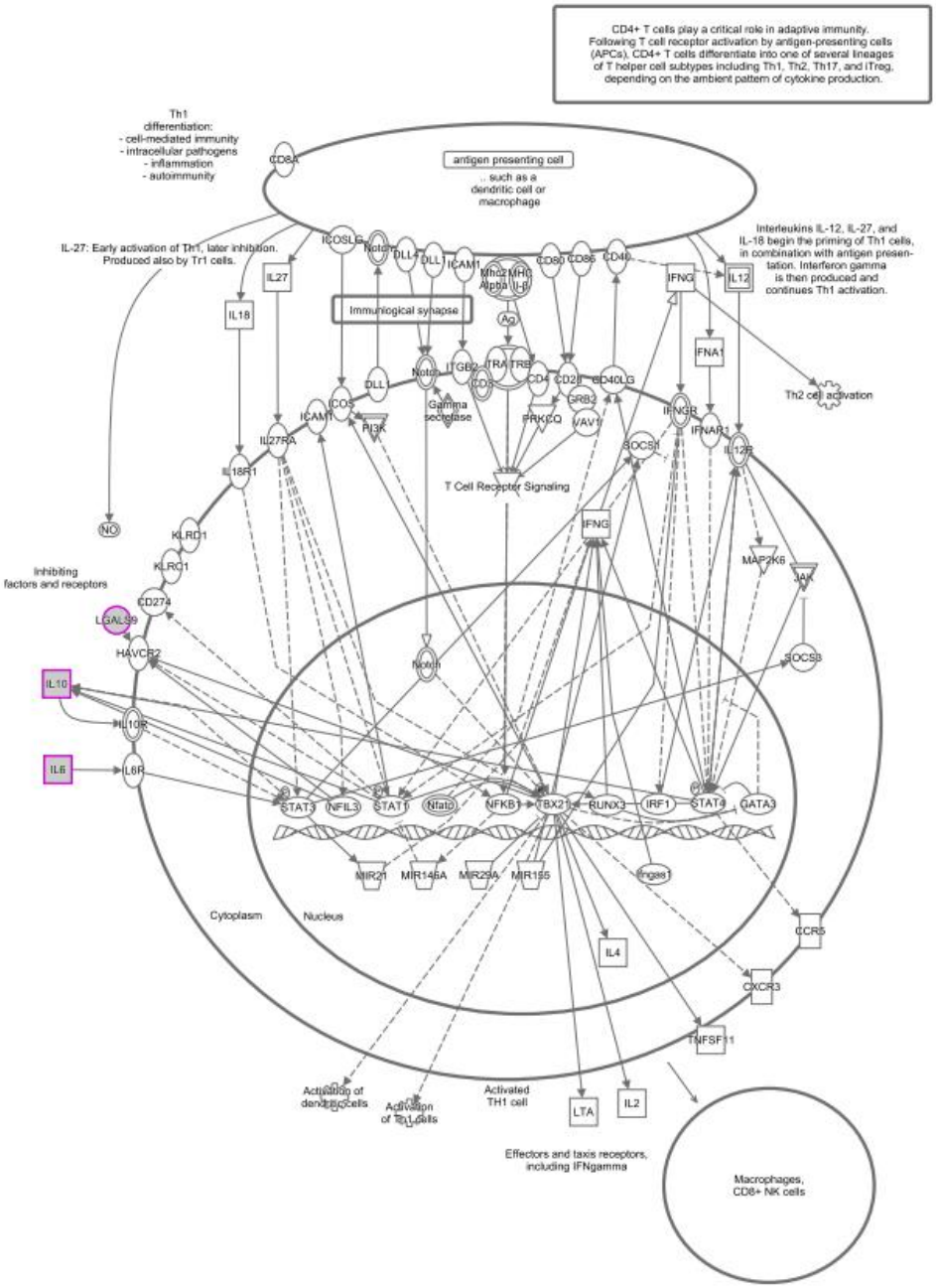


Figure 3-13. Schematic of the top canonical pathway for RV. The Th1 pathway was the top pathway for RV. The purple border highlights the candidate causal genes obtained from the signal-to-gene pipeline analysis that is implicated in the pathway: IL6, IL10 and LGALS9.

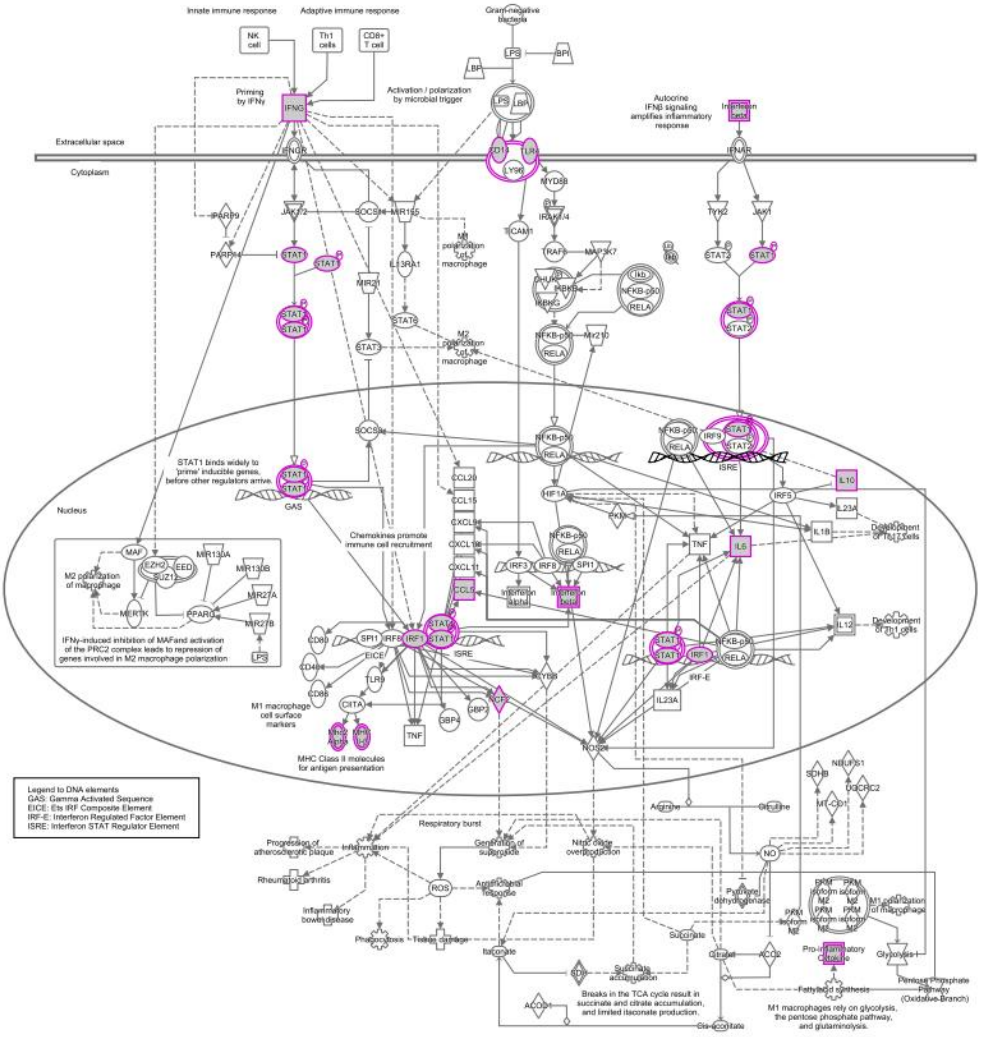


Figure 3-14. Schematic of the top canonical pathway for RSV. The macrophage activation pathway was the top pathway for RSV. The purple border highlights the candidate causal genes obtained from the signal-to-gene pipeline analysis that is implicated in the pathway.

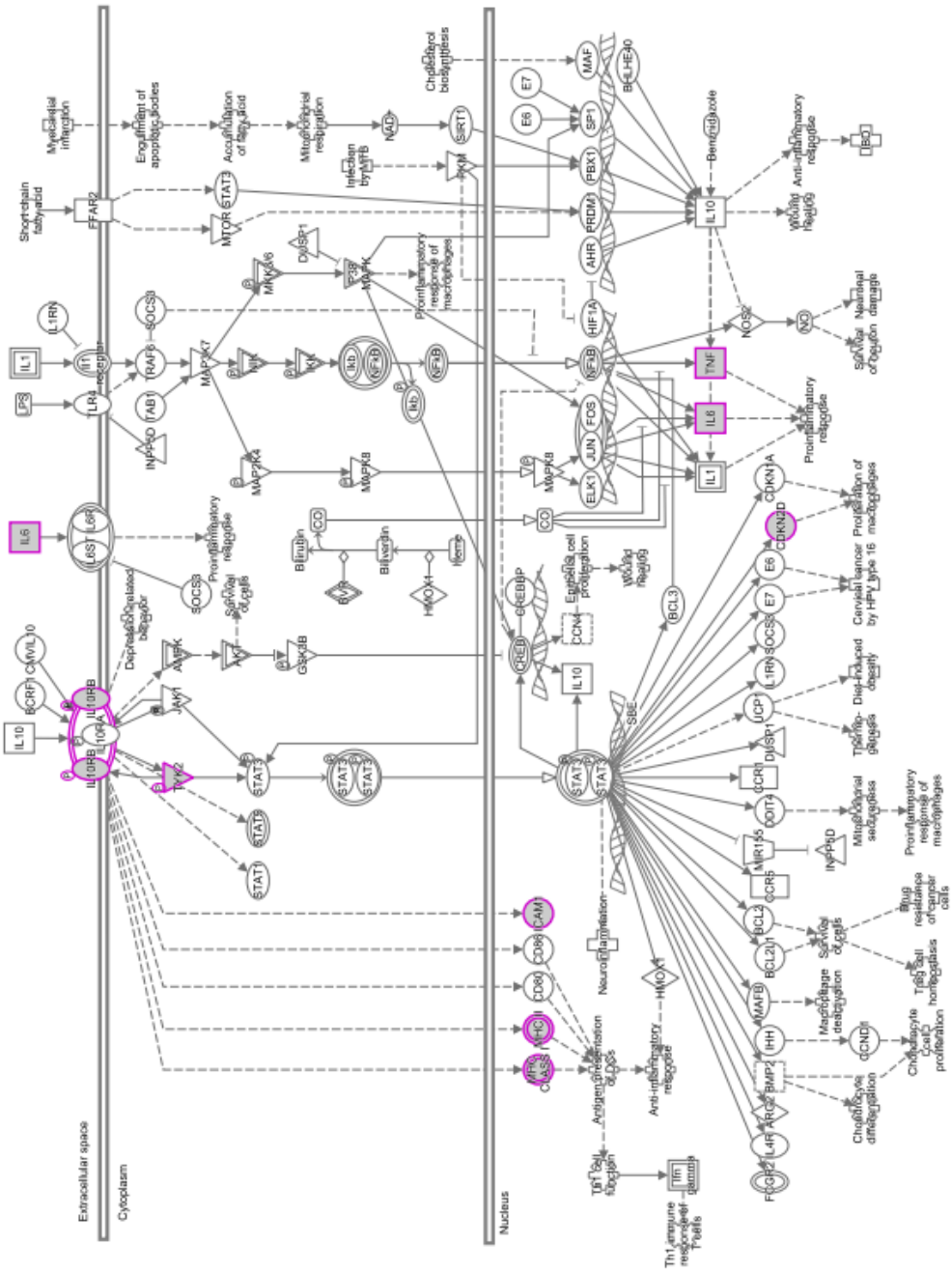
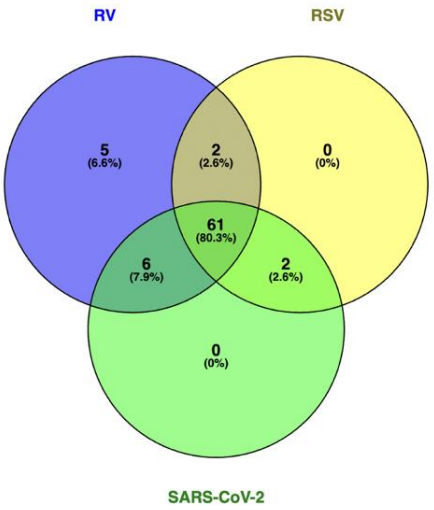


Figure 3-15. Schematic of the top canonical pathway for SARS-CoV-2. The IL-10 signalling pathway was the top pathway for SARS-CoV-2. The purple border highlights the candidate causal genes obtained from the signal-to-gene pipeline analysis that is implicated in the pathway.

3.4.5 IPA Analysis Identifies Diseases and Functions Associated with Host-Virus Interactions

Using IPA, association of diseases with host-RV, RSV and SARS-CoV-2 were identified. For the genes identified for the three viruses, a high degree of overlap (80.3%) was associated with the diseases (Figure 3-16A). 74, 65 and 69 disease pathways were identified for RV, RSV and SARS-CoV-2 respectively with 61 diseases being common for the three viruses. The top diseases identified for RV were diseases involving cellular development (Figure 3-16B). Connective tissue disorders were the top disease for RSV (Figure 3-16B) whilst infectious diseases were the top disease for SARS-CoV-2 (Figure 3-16B). The top 10 diseases associated with SARS-CoV-2 susceptibility and severity included various diseases implicating multiple organ systems (Figure 3-16B). These included connective tissue disorders, endocrine system disorders, gastrointestinal disease, and neurological disease. Overlap in diseases between the three viruses were also observed. For RV/RSV, diseases involving cellular development, cellular growth and proliferation, development and function of hematological system, and structure and development of lymphoid tissue were common in the top 10. In contrast, only gastrointestinal disease overlapped between RV and SARS-CoV-2 in the top 10. Whereas connective tissues disorders, inflammatory disease, immunological disease, inflammatory response and organismal injury and abnormalities were found to be common for both RSV and SARS-CoV-2. These results suggest that infection involving SARS-CoV-2 is more systemic and affects multiple organ systems whereas diseases implicating the immune response are important for all three viruses. Furthermore, RSV and SARS-CoV-2 had no unique diseases. In contrast, 5 diseases were found to be unique to RV. These included behavior, DNA replication, recombination and repair, organismal functions, amino acid metabolism and post-translational modifications.

A)



B)

RV	RSV	SARS-CoV-2
Cellular Development	Connective Tissue Disorders	Infectious Diseases
Cellular Growth and Proliferation	Immunological Disease	Organismal Injury and Abnormalities
Hematological System Development and Function	Inflammatory Disease	Endocrine System Disorders
Lymphoid Tissue Structure and Development	Inflammatory Response	Gastrointestinal Disease
Cell Cycle	Organismal Injury and Abnormalities	Immunological Disease
Gene Expression	Skeletal and Muscular Disorders	Metabolic Disease
Digestive System Development and Function	Cellular Development	Inflammatory Disease
Gastrointestinal Disease	Cellular Growth and Proliferation	Neurological Disease
Hepatic System Development and Function	Hematological System Development and Function	Connective Tissue Disorders
Hepatic System Disease	Lymphoid Tissue Structure and Development	Inflammatory Response

Figure 3-16. A) Venn diagram showing disease overlap identified for Host-RV/ RSV/ SARS-CoV-2 Interactions. B) Table showing the top 10 diseases for RV, RSV and SARS-CoV-2. Pathway common between RV/RSV is highlighted in green, yellow for RV/SARSCoV-2 and orange for RSV/SARS-CoV-2.

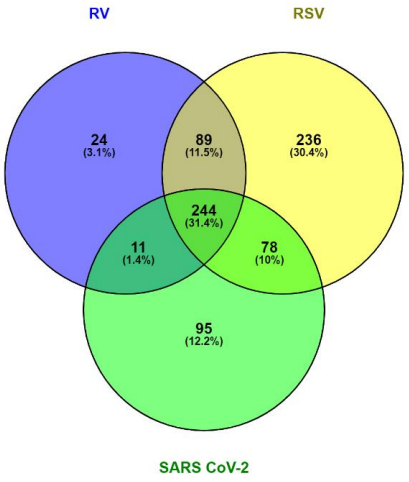
3.4.6 Upstream Regulators of Genes Identified for Host-RV/RSV/SARS-CoV-2 Interactions

Upstream analysis associated with the three viruses were further investigated using IPA. 244 upstream regulators were found to overlap between the three viruses (Figure 3-17A). The most significant upstream regulator for RV was identified as Neuropeptide Y (NPY) with IL10 and IL6 being the target molecules (Figure 3-17B). TLR4, the most significant upstream regulator for RSV, targeted CCL5, CCR5, CXCL8, IFNG, IL10, IL13, IL15, IL18, IL4, IL6, MMP1, MX1 (Figure 17B). For SARS-CoV-2, forkhead Box C1 (FOXC1) was the most significant upstream regulator. The target molecules for FOXC1 were HLA-C, HLA-DQB1, HLA-F, ICAM1, IFITM3, IFNAR1, IFNAR2, MUC1, MX1, OAS1, OAS2, OAS3 (Figure 3-17B). For RV, the top 10 upstream regulators consisted of cytokines, G-protein coupled receptors, kinase, and transmembrane receptors. In contrast, the top 10 for RSV was populated with transmembrane receptors such as TLR4, TLR2, TLR6 and TLR9. Whereas transcription regulators including FOXC1, high mobility group box 1 (HMGB1), STAT1 and twist-related protein 1 (TWIST1) showed association for SARS-CoV-2 in the top 10.

Cytokines were identified in the top 10 for all three viruses. These data suggests that the expression of other molecules may be regulated by cytokines for all three viruses. Whereas transmembrane receptors and transcriptional regulators may influence expression of other molecules for host-RSV and host-SARS-CoV-2 interactions respectively. 24, 236 and 95 upstream regulators were unique to RV, RSV and SARS-CoV-2 respectively. AKT Serine/Threonine Kinase 1 (AKT1) is one of the upstream regulators unique to RV. Whereas, TGF-β1 was unique to RSV. 6-phospho-fructo-2-kinase/fructose-2,6-biphosphate-3 (PFKFB3) was an upstream regulator unique to SARS-CoV-2. Novel upstream regulators for the three viruses were also identified. latent-transforming growth factor beta-binding protein-2 (LTBP2), lysosome-associated membrane protein 2 (LAMP2) RNA, Ro60-associated Y3 (RNY3) were unique for RV, RSV, SARS-CoV-2 respectively. Overlapping, unique and novel upstream regulators can serve as molecular

targets for further understanding host-RV, RSV and SARS-CoV-2 pathogenesis.

A)



B)

RV	RSV	SARS-CoV-2
NPY	TLR4	FOXC1
ADIPOR1	IFNG	IFNG
MERTK	TLR2	TWIST1
ADRB2	IL27	IL1RN
HLA-DQ	TLR6	HMGB1
CLEC7A	IL4	RNY3
IFNL1	TLR9	TNF
TLR5	IL10	MAPKAPK2
LILRA2	IL1B	BTK
JAK3	NFkB (complex)	STAT1

Figure 3-17. A) Venn diagram showing predicted upstream regulators overlap for Host-RV/ RSV/ SARS-CoV-2 Interactions. B) Table showing the top 10 upstream regulators for RV, RSV and SARS-CoV-2. The colours indicate types of upstream regulator with yellow:other, pink:trans membrane receptor, blue:kinase, purple:G-protein coupled receptor, grey:complex, green:cytokine, white:transcription regulator.

3.4.7 Analysis of Gene-Drug Interaction and Clinical Trial Data Identifies Genes as Druggable Targets

Candidate causal genes identified from the signal-to-gene analysis for susceptibility to each virus could be druggable targets. This can be utilised for the purpose of drug repurposing or to facilitate the development of novel therapeutics. The database “clinical trials.gov” was interrogated to identify any drugs that targeted the candidate causal genes and were currently in phase II, III and IV of clinical trials. For genes associated with RV, RSV and SARS-CoV-2 susceptibility and severity, 1 , 16 and 11 genes were targeted by drugs that were in phase II, III and IV respectively. There was a diverse range of therapeutics ranging from anti-cancer drugs to monoclonal antibodies. To explore the interaction between the candidate causal genes and drugs, the drug-gene interaction database was examined. The highest drug interaction associated with RV infection was lactose. Siltuximab and olokizumab, which targeted IL6, were the highest scoring monoclonal antibodies for RV. Interestingly, IC14 was the highest scoring monoclonal antibody for RSV with an interaction score of 92.74. Commonly used to treat sepsis, IC14 targets CD14. Elezanumab, an anti-RGMA antibody, had the highest interaction score at 61.83 for SARS-CoV-2 along with anti-CCR9 antagonist MLN3126 and vercirnon.

3.5 Discussion

The need for greater understanding of respiratory viral infections is imperative as they are one of the leading causes of mortality and morbidity worldwide. The main aim of the Chapter was to conduct a signal-to-gene/pathway analysis and drug-gene interaction analysis using genetic signals associated with RV/RSV/SARS-CoV-2 susceptibility and disease severity. Through interrogation of multiple datasets, 6, 108 and 119 candidate causal genes were identified to be relevant to RV, RSV and SARS-CoV-2, respectively. IL6, STEAP1B and IL6-AS1 were implicated as important for all three viruses. Phenotypes relating to respiratory, blood cell counts, allergic, inflammatory, and auto-immune and viral traits were identified to be associated to the genetic signals using phenome-wide association analyses. Association between these traits and the signals showed excellent agreement with the identified gene-based pathway analyses. Of note, pathways involving modulation of the immune system may be particularly important for interactions between the host and RV, RSV, and SARS-CoV-2 interactions. Analysis of disease associations showed that SARS-CoV-2 infection may impose a more systemic effect on the host with an overarching influence on multiple organ systems, as opposed to RV and RSV, which were more restricted. Insight into the site of action of the three viruses were given following expression profiling and the enrichment of the candidate causal genes in lung and blood tissues. Initial interactions between the virus and the host airway and then the subsequent inflammation is important. Analysis of drug-gene interaction highlighted opportunities for the individual viruses including confirmation of drugs currently in use for example the treatment of SARS-CoV-2 infection using anti-IL6 antibodies. The development of novel therapeutics may be aided by targeting the genes, pathways and upstream regulator identified from the analysis.

Only 24 candidate causal genes were found to overlap between the three viruses out of the 206 identified with only IL6, IL6-AS1 and STEAP1B being common for all three (Figure 3B). NF-kB-mediated IL6 production has been documented to be stimulated by RV infection[456]. Moreover, RV-induced production of IL6 may potentially be associated with pyrexia, systemic

symptomatology, infiltration of submucosal lymphocytes, and anti-viral antibody response[456]. IL6, an anti-inflammatory regulator, is thought to be important in the RSV infection disease severity. Following RSV infection, lung resident immune cells are stimulated by IL6 which produces IL27 and subsequently drives the suppression of inflammation[191]. Negative correlation between IL6 and partial pressure of oxygen in arterial blood and peripheral oxygen saturation was observed in individuals with SARS-CoV-2 infection[457]. This has been suggested to be associated with respiratory failure. Furthermore, IL6 was identified as a relevant predictor of SARS-CoV-2 disease outcome[457]. In comparison to IL6, few reports implicate IL6-AS1 and STEAP1B in RV, RSV or SARS-CoV-2 infection.

The difference in the three infections is indicated by the identification of only a small number of common genes between the three viruses. Genes with a wide variety of different functions were found to be unique for RSV and SARS-CoV-2. Genes with immunoregulatory functions such as CCR5, CCL5, IL4, IL8, IL13, IL15, IL18, IFNG and CD14, and metalloproteinases such as ADAM33, MMP1 and MMP16 were unique to RSV. Mitochondrial catalysis genes for example BCO2, COX6B1 were also found to be unique to RSV. Other unique genes to RSV included CYBA, implicated in microorganism death, ELOA, involved in elongation rate of RNA polymerase II transcription, and FAN1 with functions in DNA interstrand cross-link. In contrast, genes encoding adhesion molecules such as CDH15, ICAM1, ICAM5 were unique to SARS-CoV-2. Apolipoprotein family members APOE and APOM were also found to be unique for SARS-CoV-2. RASIP1, which regulates Rho GTPase signalling, gene-encoding the epigenetic reader protein TCF19, TNXB which stimulates collagen fibrillogenesis and WNT3 a modulator of embryogenesis was some of the genes unique to SARS-CoV-2[449–451,458]. These analyses provide a framework to further investigate the identified candidate causal genes for the individual virus and importantly the overlapping genes between two or three viruses.

Immunomodulatory pathways were identified as the top pathway for all three viruses. Asthma exacerbation induced by RV infection is postulated to stimulate inflammation of the airways through either the inhibition of virus-specific Th1 responses, the enrichment of Th2 responses, or both to alter the downstream T-cell responses[459]. For RSV identified genes, the macrophage activation pathway was the top disease. Infection of alveolar macrophages with RSV showed an increase in neutrophil infiltration and exacerbation of allergic airway inflammation[460]. Genes involved in activation of macrophages were upregulated in RSV-challenged lung macrophage cells[461]. The top pathway for SARS-CoV-2 was IL10 signalling. Pro-inflammatory cytokines such as IL10 were produced upon SARS-CoV-2 infection. IL10 production along with macrophage activation can result in vascular hemostasias which can lead to capillary leak syndrome, thrombosis and disseminated intravascular coagulation[462]. Subsequently ARDS can develop which can lead to the failure of multiple organs and death[462]. However, regulatory cells such as Tregs can produce IL10 to antagonize the overactive immune response therefore acting as a strong anti-inflammatory agent[462]. Furthermore, IL10 could present as a prognostic biomarker as significant association between IL10 production and COVID-19 disease severity has been documented[463].

Pathways unique to RV, RSV and SARS-CoV-2 were also identified. The Th17 activation pathway was found to be unique to RV. The IL-23/Th17 pathway is an important inflammatory pathway which mediates abnormal inflammatory responses[464]. A study by Teng *et al* found that children with RV infection had significantly elevated levels of Th17 and reported that IL-23/Th17 contributes to the development of RV-associated wheezing[464]. Whereas neutrophil extracellular trap (NET) signaling pathway and RAR activation were some of the unique pathways for RSV. Neutrophil extracellular traps are DNA structures decorated with cytosolic, nuclear, and granular proteins which entrap pathogens.[465] During RSV infection, neutrophils release NETs to prevent RSV from binding to host cells thereby limiting infection[466]. In contrast, little literature exists on RAR activation and RSV infection. Retinoic acid receptor (RAR), which serves as a ligand-activated transcription factor, is activated by all-trans retinoic acid and 9-cis-retinoic acid[467]. Therefore further investigation may be

useful. For SARS-CoV-2, pathways found to be unique included coronavirus replication pathway and IL-22 signalling. Although coronavirus pathogenesis was found to overlap between the three viruses, coronavirus replication was unique to SARS-CoV-2. Whereas Il-22, a member of the IL-10 family of cytokines, was reported to be elevated in hospitalised COVID-19 patients and associated with cardiac risk during hospitalisation[468]. Moreover, novel pathways implicating autoimmune disorders such systemic lupus erythematosus and rheumatoid arthritis were observed between the three viruses. These include systemic lupus erythematosus In T cell signaling pathway and systemic lupus erythematosus In B cell signaling pathways. The genes and signalling molecules involved in overlapping or unique pathways may serve as useful targets for therapeutic approaches to treat respiratory viral infections.

Diseases involving cellular development, connective tissue disorders, and infectious diseases were the top disease associations identified with RV, RSV and SARS-CoV-2, respectively. In comparison to RV and RSV, diseases affecting various different organ systems such the endocrine, gastrointestinal and nervous system were identified for SARS-CoV-2. Only five disease classes were unique to RSV whereas RSV and SARS-CoV-2 had zero unique disease classes. Several overlapping diseases such as endocrine system disorders were identified. Little is known about the influence of RV infection on the endocrine system. With regards to RSV, a retrospective study showed that low sodium levels (hyponatraemia) were common among a cohort of 130 infants with RSV infection[469]. Furthermore, four infants suffered hyponatraemic seizures[469]. As one of the most frequent comorbidities associated with COVID-19 severity, the endocrine disorder diabetes and its relationship with COVID-19 is a topic of much interest. SARS-CoV-2 may inflict damage by binding to the abundantly expressed ACE2 on the pancreatic β-cells[470]. Severe complications of pre-existing diabetes such as diabetic ketoacidosis and new onset of diabetes in individuals with COVID-19 have also been reported[471,472]. Furthermore, COVID-19 susceptibility and severity may be impacted by obesity, adrenal dysfunctions and hypovitaminosis[470]. Gastrointestinal disease was one of the overlapping diseases for RV, RSV and SARS-CoV-2. Infection

of the digestive tract by SARS-CoV-2 has also been documented with the GI tract showing high expression of the ACE2 receptor[473]. The increase in COVID-19 incidence could be attributed to the effect of elevated blood glucose levels on ACE2 expression. In comparison, there is little literature on RV/RSV infection and gastrointestinal disease.

Neurological disease was one of the diseases showing overlap between the three viruses. A case report was reported of a previously healthy boy who acquired RV-associated acute encephalitis[474]. Neurological symptoms such as seizures and disturbance of the consciousness were observed[474]. A meta-analysis assessing the neurological complications of RSV infection reported that seizure was the most commonly reported neurological manifestation[475]. Moreover, the central nervous system of 12 cases was positive for RSV detection[475]. The initial report from Wuhan China highlighted that 36.4% of patients with COVID-19 exhibited neurological symptoms[476]. These manifestations could be characterised into three categories: central nervous manifestations (headache, dizziness impaired consciousness, seizure etc.), peripheral nervous system (impairment of smell, taste and vision, and nerve pain) and skeletal muscular injury[476]. Furthermore, 73% of hospitalised COVID-19 patients from a cohort of 2533 patients exhibited neurological symptoms[477]. Another overlapping disease between the three viruses was connective tissue disorder. However, little has been reported about connective tissue disorder and its association with RSV or RV infection. In contrast, risk of autoimmune and autoinflammatory connective tissue disorders has been associated with COVID-19[478]. In a retrospective study, association between COVID-19 severity and the following autoimmune and connective tissue disorders such as rheumatoid arthritis, ankylosing spondylitis, Crohn's disease, ulcerative colitis, psoriasis, and sarcoidosis were identified[478].

Exploration of upstream regulators revealed that NPY, TLR4 and FOXC1 were the top upstream regulators for RV, RSV and SARS-CoV-2 respectively. NPY helps mediate numerous physiological processes in the immune, cardiorespiratory, endocrine and the nervous system[479]. As a neuromodulator NPY facilitates the release of neurotransmitters such as dopamine[480]. NPY

also impacts different airway functions and helps regulate the tone of the airways and the smooth muscle[479]. During asthma exacerbation, the levels of NPY in serum are elevated[479]. TLR4 is an innate immune receptor which forms a complex with MD-2 to detect polysaccharide, a gram-negative bacterial surface component[251]. In response to viral infection, TLR4-mediated signalling also stimulates the production of Type I interferons and proinflammatory cytokines[251]. Interaction between RSV fusion protein and RSV through MD-2 has also been documented[481]. Furthermore, RSV-induced TLR4 signalling has shown to be inhibited by the TLR4 antagonists *Rhodobacter sphaeroides* LPS and synthetic E5564[482]. FOXC1, believed to have a role in embryonic development, is relatively unknown in regard to pathways and mechanisms[483]. However, pathogenic variants of FOXC1 are associated with Axenfeld-Rieger syndrome (ARS) and anterior segment dysgenesis[483]. ARS is characterised by distinctive facial features, and skeletal, dental, cardiac and audiologic abnormalities[483].

IPA analysis also identified upstream regulators unique to the three viruses. One of the upstream regulators unique to RV is protein kinase B (AKT1). Upon exposure to external stimuli, PI3K is activated resulting in the production of PIP3[484]. Recruitment of AKT to the membrane is then initiated[484]. Through the PH binding domain, AKT then interacts with PIP3 to exert downstream effects such cell migration, proliferation, cell growth and survival and angiogenesis manning[484]. With regards to RV, a study has shown that PI3K and subsequently AKT is activated upon RV infection[485]. Furthermore, AKT activation is required for maximal IL8 expression, a key cytokine suggested to be implicated in RV-induced asthma and COPD exacerbations[485]. TGF-β1, a unique upstream regulator to RSV, is a key driver of developmental and cell behaviour processes[486]. Gene expression is regulated by TGF-β through receptor-mediated activation of SMAD transcription factors[486]. RSV infection has shown to enhance RSV replication by inducing cell cycle arrest[487]. Moreover, application of TGF-β1 signalling inhibitors and anti-TGF-β1 antibody led to the rescue of RSV-induced cell cycle arrest[487]. PFKFB3 was identified as a unique upstream regulator for SARS-CoV-2. A study explored the mechanisms underlying dysregulated glucose metabolism and how this

could facilitate development of a severe SARS-CoV-2 infection in diabetic patents[488]. Glycolytic flux was shown to impact SARS-CoV-2 infection[488]. Replication of SARS-CoV-2 and cytokine response was found to be abolished following inhibition of PFKFB3 and lactate dehydrogenase A[488]. This suggests that glycolysis could be an important upstream event during SARS-CoV-2 pathogenesis[488].

Novel upstream regulators associated with RV-RSV and SARS-CoV-2 susceptibility and severity have been identified. A major component of lung and arterial tissue, and ciliary zonule, LTBP2 is a secreted extracellular matrix protein[489]. LTBP2 has also been suggested to be a biomarker for idiopathic pulmonary fibrosis and connective tissue disease-associated interstitial lung disease[489]. LAMP2, identified to be a novel upstream regulator for RSV, is a highly glycosylated lysosomal membrane protein with essential function in chaperone mediated autophagy[490]. LAMP2 mutation is associated with a triad of cardiomyopathy, myopathy and encephalopathy presentation in individuals with the X-linked Danon's disease[490]. This skeletal and cardiac muscular pathogenesis is attributed to inefficient biogenesis and maturation of lysosomes[490]. RNY3, a novel upstream regulator for SARS-CoV-2, is a part of the Y-RNA family[491]. They have been proposed to be implicated in RNA-protein complex scaffolding and assembly[491]. Moreover, Y-RNAs are differentially expressed in individuals with allergic asthma and dysregulated in various different cancers[492]. A study investigated the effect of RNY3 in Th2 cell inflammatory response of allergic asthma[492]. It was reported that Jurkat cells transfected with RNY3 led to a significant increase in the mRNA levels of IL13 and decrease in cell proliferation[492]. Identification of key upstream regulators and their relationship towards downstream genes in pathways is useful for gaining a more mechanistic understanding and for deciphering the molecular basis of the viral infections. Furthermore, targeting these unique upstream regulators may potentially be useful when specific approaches to target specific viruses when developing new drugs.

Drug targets with mechanisms supported by human genetic evidence have shown positive impact in successful development[493]. In a retrospective study, progression for genetically supported drug targets through the drug development were twice as likely compared to targets with genetic support[390]. Investigation of therapeutics targeting the candidate casual genes for the treatment of several diseases have been reported in clinical trials. For example, the efficacy of anti-IL6 monoclonal antibody Clazakimumab for the treatment of active psoriatic arthritis was investigated in a randomized, double-blind clinical trial[494]. Currently the treatment of COVID-19 with Clazakimumab are being investigated in clinical trials[495]. A randomized double-blind clinical trial investigated the efficacy of Clazakimumab in improving patient outcomes in patients hospitalised for severe or critical COVID-19 with hyperinflammation[495] . Drug-gene interactions should be carefully interpreted for the purpose of drug repurposing. For example, many of the drugs targeting the candidate causal genes are chemotherapy drugs which may not be suited for repurposing. Therapeutic opportunities to treat respiratory viral infections remain limited despite being one of the leading causes of mortality. Currently, no licensed vaccines for RV exist[496]. Therefore, it is important to develop new therapeutics or repurpose drugs to treat respiratory viral infections particularly in susceptible individuals.

In summary, a significant shift in our understanding of genes and pathways have been provided by these findings which may be particularly relevant for determining the susceptibility and severity of three major respiratory viruses. These novel findings provide a significant opportunity for the development of new therapeutic approaches which could target the effects of respiratory viruses. By using this genetics-based approach and studying the similarities and differences between the viruses, this can lead to effective drug-repurposing for broad anti-viral activity.

4.0 Using Transcriptomics to Study How Primary Human Epithelial Cells Respond to Prolonged Virus Infections

4.1 Introduction

In the last Chapter I set out to leverage genetic association studies to identify genes and pathways that may be important in the host response to three common respiratory viruses. In the current Chapter I aimed to leverage transcriptomic data to characterise the cellular response to these same viruses with scope to provide new insight but importantly provide support for the role of genes and pathways that I identified in Chapter 3. Transcriptomic profiling is a vital part of genomic research and meaningful biological insights have been garnered through the use of RNA-seq[497]. The technique offers a framework for exploring differential gene expression, novel transcripts, allele-specific expression, and alternative splicing[498].

In transcriptomic sequencing, RNA is extracted, and a library preparation protocol is prepared which either enriches or depletes a total RNA sample for a specific RNA species[357]. This allows the removal of transcripts that consumes the bulk of sequencing reads so that less-abundant RNAs can be detected, and the overall depth of sequence coverage can be maintained[357,499]. For example, RNAs with polyadenylated (poly-A) can be selected for by using poly-T oligos bonded to a substrate[357]. A library of cDNA fragments is then created, and sequencing adaptors are ligated to one or both ends[356]. To retain information of the strands, which is vital for *de novo* transcript discovery, chemical labels such as deoxy-UTP are incorporated to allow the second strand cDNA to be distinguished from the first[500]. Each molecule is sequenced in a high-throughput manner to attain short sequences from either both end (paired-end sequencing) or one end (single-end sequencing)[356]. The resulting reads are either aligned to a reference genome or transcript or assembled *de novo* without a genomic sequence[356]. The reads are then

classified into three types (junction reads, exonic reads, and poly(A) end-reads) to create a base-resolution expression for each gene[356].

One of the most powerful uses of RNA-seq technology is investigating the differences in gene expression between two or more conditions through a process called differential expression. Statistical modelling via normalisation is applied during differential expression analysis to reduce variability and remove systematic biases[498]. Profiling changes in host gene expression during viral infection can give insight into the complex interplay of pathways involved in host-virus interactions. Furthermore, pathways impacted by the virus can be inferred by identifying upregulated and downregulated genes. Therefore, the use of RNA-Seq has contributed to significant discoveries in the field of respiratory medicine. A study by Yick *et al* utilised RNA-Seq to understand the molecular and cellular pathways of asthma[501]. Forty-six genes, including solute carrier family 26 member 4 (SLC26A4) and periostin osteoblast-specific factor (POSTN), were found to be differentially expressed between asthma and control[501]. SLC26A4 has shown to be associated with asthma pathophysiology whereas POSTN is highly induced in airway epithelial cells[501]. Furthermore, several of the genes were found to be involved in mRNA degradation and translational processes such as Staufen double-stranded RNA binding protein 2 (STAU2) and tryptophanyl-tRNA synthetase (WARS)[501]. The differentially expressed genes were also associated with gene networks implicated in cellular movement, morphology and apoptosis[501]. Whereas the large-airway epithelium, peripheral blood, and alveolar macrophages were profiled in a study by Morrow *et al* to identify tissue-specific gene expression and pathway signatures of COPD[502]. The smoking-related genes cytochrome P450 1B1 (CYP1B1) and aryl hydrocarbon receptor repressor (AHRR) were among the top differentially expressed genes in large-airway epithelium[502]. A significant overlap of smoking and airway disease phenotypes were found between large airway epithelium and macrophages[502]. In addition, pathways involving haemostasis and immune signalling overlapped between the three tissues for emphysema and pathways relating to amyloid and telomere for smoking[502].

Our understanding of host-viruses interactions has been strengthened by the use of transcriptomics. A study by Nicolas de Lamballerie *et al* characterised the cellular transcriptomic signatures of nasal airway epithelium following respiratory virus infection[499]. A global virogenomic signature of downregulated cilium-related gene expression, particularly those involved in ciliary dynein machinery and intraflagellar transport, was found to be shared by influenza viruses, hRSV, and hMPV[499]. Since primary respiratory infection can contribute to the establishment of secondary infection through impairment of mucociliary clearance (MCC), the interplay between the viruses and the cilium-related machinery is important especially in the context of diseases associated with defective MCC such as cystic fibrosis[499]. Whereas Chappin *et al* profiled the airway and blood monocyte transcriptome in patients with severe RSV infection[503]. Elevated expression of interferon-responsive and anti-viral genes were found to extend from the airways to the peripheral blood[503]. Furthermore, genes related to inflammation and migration were upregulated in airway monocytes[503]. In contrast, airway gene expression profile was identified by Wang *et al* using nasal brush samples from infants with RSV infection[504]. Among the differentially expressed genes were cytokines (TNF superfamily member 10 TNFSF10, IL6), extracellular matrix proteins (matrix metallopeptidase 19 MMP19, vimentin), regulators of inflammation (cluster of differentiation CD163, C-X-C motif chemokine ligand 2 CXCL2), and components of signalling processes[504]. Novel genes such as ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (ST3GAL1) and ATPase phospholipid transporting 10B (ATP10B) were also identified[504].

Several research has focused on acute RSV and RV infection. Acute viral infection is characterised by a rapid disease onset which can be resolved by robust innate immune response[505]. Mild symptoms such as sore throat, coughing, rhinorrhoea, and sneezing can manifest as early as 10-16 days[506]. The symptomatic phase peaks around 2-3 days post-infection[506]. Symptoms typically last for approximately one week with 25% of cases lasting longer[506]. Although acute RSV and RV infections are well documented, more investigation into the host response to prolonged infections is needed[507]. By studying prolonged infections, host-virus dynamics and the differences and

similarities in genes and pathways between acute and prolonged infection can be determined. Understanding the molecular basis of host-virus interactions during acute and prolonged infection can facilitate the development of therapeutic interventions that can target acute and prolonged infections.

4.2 Aims and Objectives

Overall Aim: To investigate how primary human bronchial epithelial cells respond to prolonged infection of RSV-A2 and RV-A16 using transcriptomics.

The objectives of the Chapter are:

- To identify differentially expressed genes in response to prolonged RSV-A2 and RV-A16 infection using RNA-Seq.
- To identify and compare pathways that are upregulated and downregulated between RSV-A2 and RV-A16.
- To identify and compare upstream regulators that are activated and inhibited between RSV-A2 and RV-A16.
- To investigate whether viral gene signatures provide support for genes identified from previous signal-to-gene mapping analysis.

4.3 Methods and Materials

4.3.1 Viral Infection and RNA isolation

HBECs D195307 were infected with RSV-A2 (MOI of 10) and RV-A16 (MOI of 10) as outlined on Chapter 2.3.3. RSV-A2 was purchased from Public Health England (Catalogue code: 0112092v). RSV-A2 was kindly provided by Professor Colin Bingle[369]. The donor information is detailed in table 1. HBECs were cultured in Lonza BEGM® Bronchial Epithelial Cell Growth Medium BulletKit® and were infected with the virus for 1 hour. Virus-media was removed and replaced with fresh media. RNA was isolated on day 6 (Chapter 2.1.2) and quantified using NanoDrop (Chapter 2.1.3). Biological and technical replicates were conducted in triplicates.

Table 4-1. Specifications for HBEC donor D195307 used for cell culture experiments.

Test	Minimum	Maximum	Results
Tissue Acquisition Number			19665
Donor Screen Information			
Age			19 Y
Sex			MALE
Race			CAUCASIAN
Cell Type			NHBE
QC Evaluation Medium			BEGM
Cell Strain Calculations			
Cell Passage			1
Viability-Tryp Blue Exclusion	>= 70%		91%
Cell Count (cells/amp)	>= 500,000		897,000
Total Population Doublings	>= 15		15
Seeding Efficiency	>= 20%		66%
Doubling Time (hours)	12	48	15
Sterility Test			Negative
Mycoplasms			Negative
Virus Testing			
HIV Test			Not Detected
HBV Test			Not Detected
HCV Test			Not Detected

4.3.2 RNA-Seq Methods

4.3.2.1 RNA Integrity Number

The RNA integrity number (RIN) score was calculated for each sample prior to RNA-Seq. RIN score analysis was conducted by the Deep Seq, Next Generation Sequencing Facility, University of Nottingham Medical School using the Agilent Technologies 4200 TapeStation. A numerical system on a scale form 1-10 is used to determine the RNA quality[508]. A score is assigned with 10 denoting an intact, high-quality RNA and 1 representing significantly degraded sample[509].

4.3.2.2 Library Preparation and RNA-Seq

Library preparation and RNA-Seq was performed by Oxford Genomics Centre through polyA RNA sequencing using Illumina v1.5 chemistry and the NovaSeq6000 reagent kits. Adapters, single-stranded RNA sequences ligated to sRNAs, are attached to the ends of the DNA template[510]. This immobilises the cDNA to flow cells for sequencing[510]. Sequencing primers anneal to the adapters and PCR amplification of adapter-ligated DNA fragments occur[510]. Reads are produced by a reversible-terminator cyclic reaction which is associated with a base-specific colorimetric signals[511]. The colorimetric signals are translated to a base call in a FASTQ format[511]. Each nucleotide corresponds to an ASCII-encoded quality number which is associated with a Phred score[511]. The score is an identifier for the quality of sequencing as it measures the accuracy of the base call[511].

4.3.2.3 Processing of RNA-Seq Data

RNA-Seq raw sequence reads were provided in a FASTQ format by Oxford Genomics Centre. The identifiers, nucleotide sequences, and the Phred score for each nucleotide is stored for each sequencing read in the FASTQ file. The Linux operating system was used to process the RNA-Seq data. RNA-Seq pipeline by Rajasekar *et al* was used to process the RNA-seq data[512].

4.3.2.4 Trimming and Alignment

First, the FASTQ files from multiple lanes of sequencing for each sample were concatenated. Adapter sequences and poor-quality sequences were trimmed from the FASTQ files using Trimmomatic. Adapter sequences need to be removed from the analysis since they are not part of the biological sequences and can lead to misinterpretation of data[513]. Files were aligned to the Genome Reference Consortium Human Build 37 (GRCh37) genome using the STAR software to generate Binary alignment map (BAM) files and count files for analysis in RStudio. STAR software was used as it outperforms other aligners by a factor of >50 in mapping speed and aligns to the human genome 550 million 2 x 76 bp paired-end reads per hour whilst possessing improved alignment precision and sensitivity[514]. BAM files are compressed binary versions of the sequence alignment map (SAM) format.

4.3.2.5 Analysis of Differential Expression using EdgeR and Limma

The individual count files were combined in a matrix to create a count matrix file. This was imported into RStudio version 4.1.0 where EdgeR version 4.0.16 and limma version 3.58.1 were used to analyse differentially expressed genes between different samples and conditions. Genes were annotated using the EnsDb.Hsapiens.v75 version 2.99.0 package from Bioconductor.

4.3.2.6 Visualisation of RNA-Seq Analysis

Volcano plots were generated using the ggplot2 version 4.0.2 and ggrepel version 0.9.3 packages in RStudio. Venn diagrams were produced using Venny version 2.1.0[515] to visualise overlapping genes and pathways between different conditions.

4.3.2.7 Downstream Analysis of RNA-Seq Data Using IPA

Pathways and upstream regulators were determined using Ingenuity Pathway Analysis (IPA)[407]. RNA-seq datasets for RSV-A2 and RV-A16 with the whole gene expression, p-values and log fold changes (logFC) were uploaded onto IPA. A logFC cut-off of <-2.0 and >2.0 were used for the IPA analysis. Z-scores

for pathways and upstream regulators were calculated by IPA to determine the directionality for example positive a-score indicated an increase in activity and a negative z-score indicated a decrease in activity.

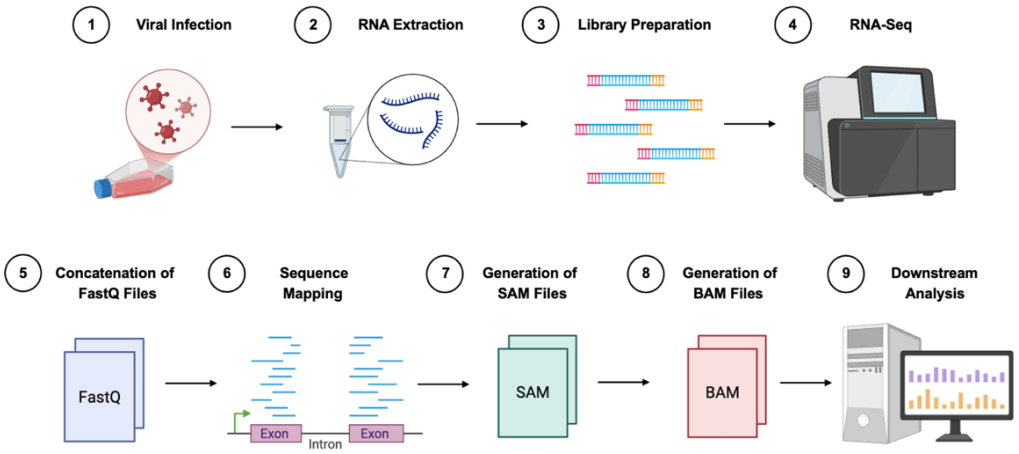


Figure 4-1. Schematic of the RNA-Seq workflow. Following infection of submerged HBECs with RSV-A2 and RV-A16, RNA was extracted, and a sequencing library was prepared for RNA-Seq. Raw sequence reads were provided in a FASTQ format. The FASTQ files were concatenated, with adapter and poor-quality sequences being trimmed using Trimmomatic. Sequencing reads were mapped to reference GRCh37 genome to generate SAM and BAM files. Count matrix file was imported into RStudio and analysis on differentially expressed genes were conducted using EdgeR and limma. IPA was used to conducted to identify pathways and upstream regulators.

4.4 Results

4.4.1 How Do Primary Human Epithelial Cells Respond To RSV-A2 and RVA-16 In Prolonged Infections?

To investigate the effect of prolonged viral infection on gene expression, HBECs D195307 grown in submerged condition were infected with RSV-A2 and RVA-16 for 6 days. RNA was extracted and taken for RNA-Sequencing.

4.4.1.1 Quality Control of RNA-Seq Data

Raw counts were transformed into counts per million (CPM) to normalise RNA-Seq data for sequencing depth (the total number of sequenced reads)[508]. The cpm function in EdgeR was used to convert raw counts into CPM and log-CPM. This transformation accounts for differences in library sizes. The mean sequencing depth was 23.47878 and the median sequencing depth was 30.95317. For downstream analysis, genes with low or no expression were filtered out from across all the samples. The filtrByExpr function in EdgeR was to filter out any lowly or non-expressed genes. Removal of low count genes also enables the estimation of the mean-variance relationship with greater reliability. In addition it reduces the number of statistical tests that need to be conducted. Figure 4-2 shows the density of log-CPM values for raw data and filtered data. The log-CPM threshold value, which is equivalent to a CPM value of approximately 0.2, is denoted by a dotted vertical line.

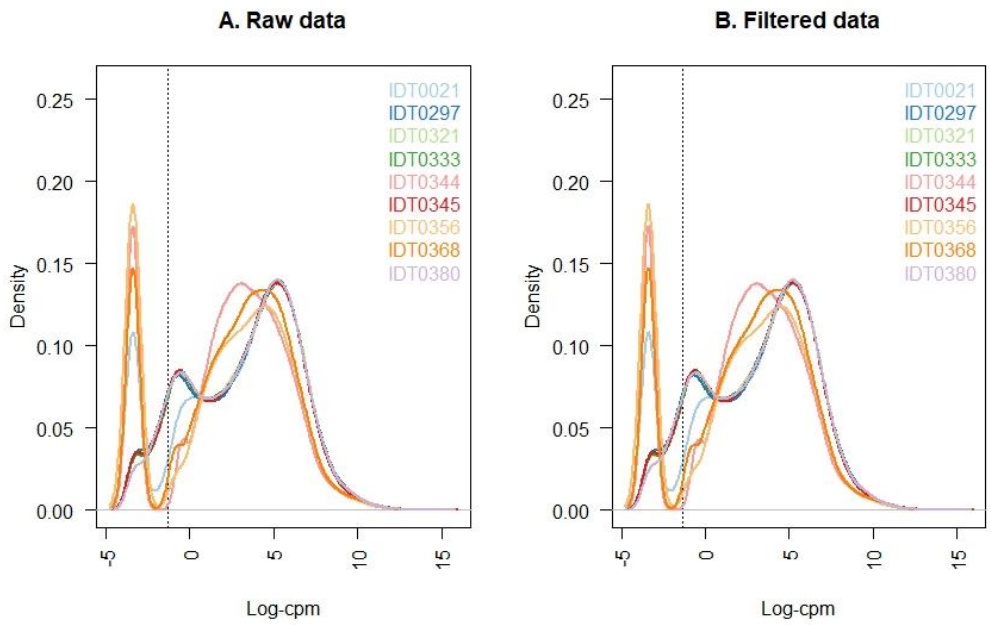


Figure 4-2. Raw data vs filtered data. Genes with little or no expression have been filtered out. Log-CPM values for raw and filtered data are displayed on the graphs. The dotted vertical line is the log-CPM threshold value which is equivalent to a CPM value of 0.2.

External factors that are not of biological interest can influence the expression of individual samples during sample or sequencing process[516]. Therefore it is important to normalise the dataset to ensure that the distribution of expression is similar across the different experiments. The calcNormFactors function in EdgeR was used to perform the method of trimmed mean of M-values (TMM) to normalise the data. The TMM method estimates the relative RNA production levels from RNA-seq data and normalises the raw read counts by trimming away extreme log fold changes of the data[517]. Figure 4-3 shows the expression distribution for unnormalised and normalised data.

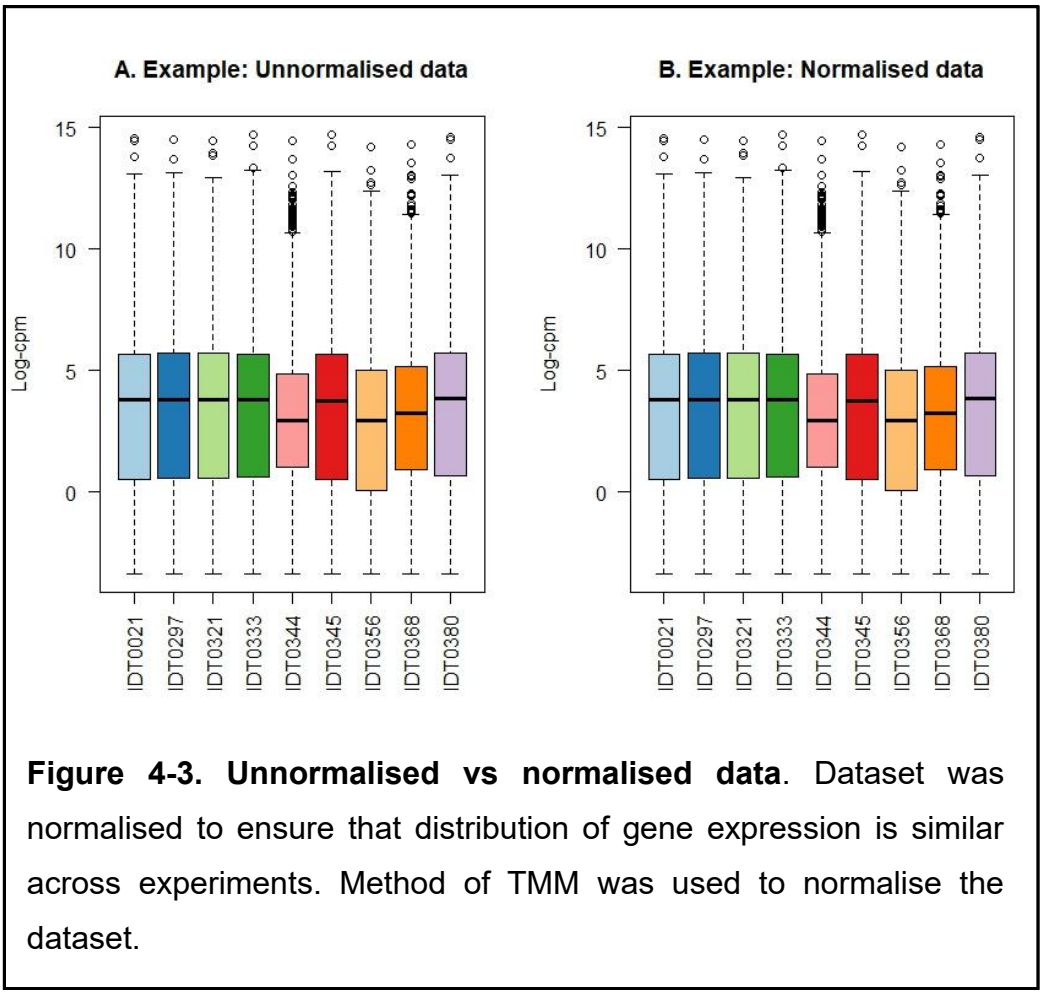


Figure 4-3. Unnormalised vs normalised data. Dataset was normalised to ensure that distribution of gene expression is similar across experiments. Method of TMM was used to normalise the dataset.

Clustering of samples though multi-dimensional scaling (MDS) plots gives an insight into the extent of differential expression[516]. The plotMDS function of limma was used to cluster the samples. Dimension 1 represents the leading fold change (leading logFC) that separates the samples[516]. Leading logFC is the average log_2-fold change for the 500 genes which are most divergent between each pair of samples by default[516]. Figure 4-4 shows MDS plots of log-CPM values between the samples. Samples are clustered based on the experimental conditions: control, RVA-16 and RSV-A2. However, RSV-A2 experiments 3 and 2 cluster more closely to each other compared to E1.

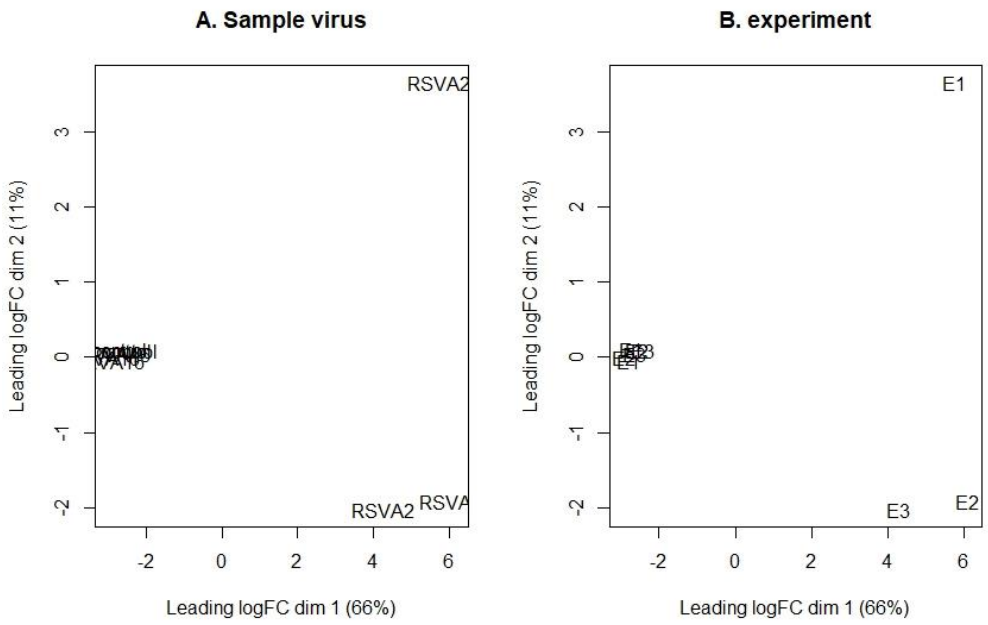


Figure 4-4. Clustering of samples using MDS plots. Samples were clustered based on experimental conditions. MDS plots of log-CPM values are displayed. Biological and technical replicates were conducted in triplicates.

Differential expression analysis was conducted using limma. Linear modelling was fitted to log-CPM values to extract residual variances[516]. The mean variance relationship of the log-counts was estimated using precision weights calculated by the voom function[516]. Raw counts was converted to log-CPM values through voom by automatically extracting library sizes and normalisation factors. Figure 4-5A shows the mean (x-axis) and variance (y-axis) of each gene and the dependence between the two before voom is applied. A decreasing trend is shown between the means and variances. This can be attributed to technical or biological variation. Sharp decreasing trends are typically observed in experiments with low biological variation whereas flatter trends are seen in experiments with high biological variation. This suggests that the experiments have low biological variation. The imFit ant contrasts.fit functions of limma is used to conduct the linear modelling. A separate model is fitted onto the expression values of each gene by the function. Limma Bayes empirical analysis was carried out by obtaining information across all genes to identify more precise estimates of gene-wise variability[516]. voomWithQualityWeights function was used to incorporate sample-level weights together with abundance dependent weights that were estimated by voom. The residual variances of the model was plotted against the mean expression values. Figure 4-5B shows the removal of the trend therefore indicating that variance is no longer dependent on the mean expression level[516]. The horizontal blue line represents the residual standard deviation mean log-CPM values[516].

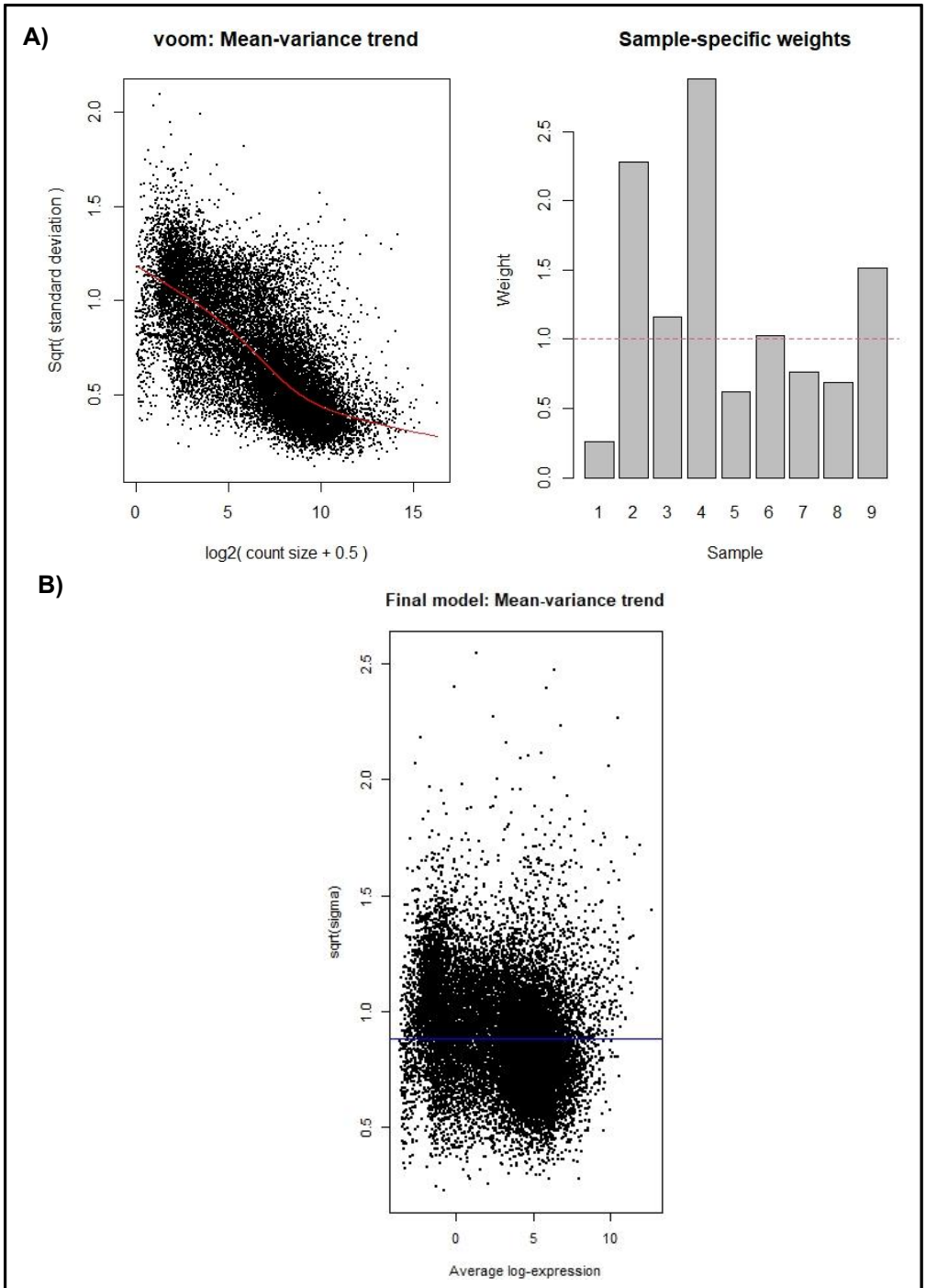


Figure 4-5. Application of precision weights by voom to estimate the mean-variance trend. Residual variances were removed by fitting linear modelling onto the log-CPM values. Precision weights were calculated by voom to estimate the mean-variance trend.

Quality weights function was applied to improve power for detecting known changes in gene expression[498]. Here the sample-level weights are incorporated with the abundance dependent weights estimated by voom using a log-linear model[498]. The approach also generates fewer false positives[498]. Figure 4-6 shows that p-values are more significant after quality weights were applied.

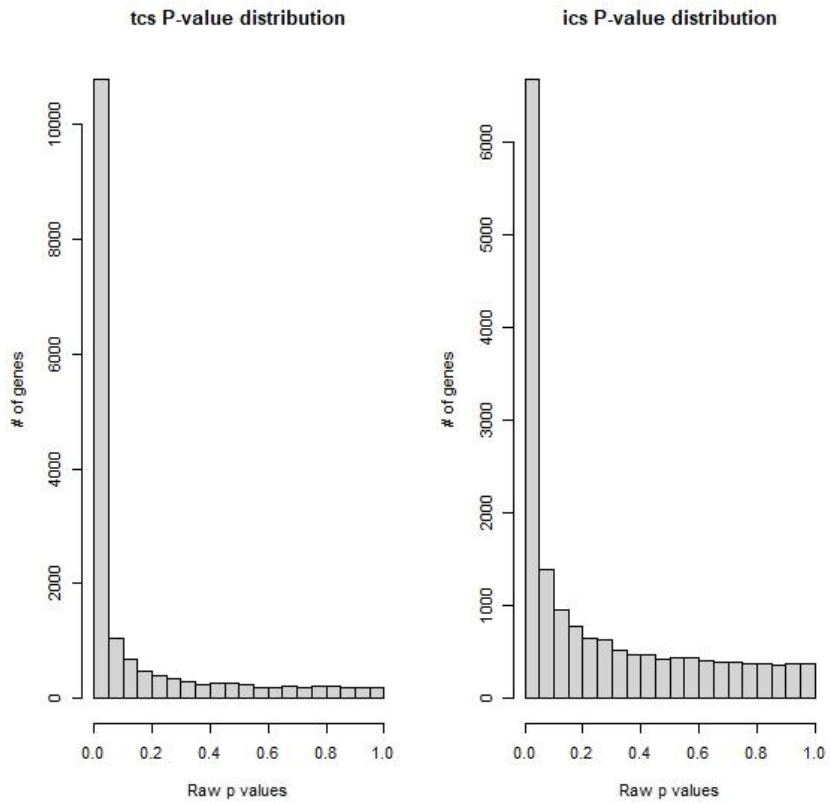


Figure 4-6. P-value distribution after application of quality weights. The abundance dependent weights estimated by voom was incorporated with sample-level weights using a log-linear model.

4.4.1.2 Differential Expression of Genes in Response to Prolonged RSV-A2 Infection

Infection of HBECs with RSV-A2 for 6 days resulted in 11190 differentially expressed genes (5% FDR). 5059 genes were upregulated whereas 6131 genes were downregulated. Figure 4-7 shows a volcano plot of the differentially expressed genes with the top 20 most significant genes for RSV-A2 with a fold change of 2 being labelled. Genes with lower p-values are higher up on the plot. Statistically significant genes that did not reach logFC > 0.5 are black whereas non-significant genes are grey. The top upregulated genes include X inactive specific transcript (XIST) and neurofilament medium chain (NEFM). XIST is involved in the X activation developmental process[518]. NEFM encodes for the medium neurofilament protein which is involved in the maintenance of the neuronal cytoskeleton[518]. Genes involved in DNA binding such as zic family member 5 (ZIC5), homeobox D13 (HOXD13) and homeobox C9 (HOXC9)[518]. RNA binding gene, Lin-28 homolog B (LIN28B), and ATP binding cassette subfamily B member 1 (ABCB1) was also among the top 20 most significant genes. ABCB1 encodes for ATP-binding cassette transporters which are involved in transporter activity[519]. In contrast, a number of keratin-encoding genes such as keratin 17 (KRT17), KRT6A, KRT5, KRT16, KRT6B, are among the top downregulated genes[518]. The top genes were ranked based on logFC as it prioritises genes with greater biological magnitude of change. This spotlights genes that are biologically relevant which can facilitate downstream analysis to identify pathways, upstream regulators and biomarkers. The list of top 10 upregulated genes by logFC for RSV-A2 is shown on table 4-2 and the top 10 downregulated genes by logFC on table 4-3. Other top downregulated genes include transporter cell adhesion gene collagen type XVII alpha 1 Chain (COL17A1) transmembrane activity gene solute carrier family 38 member 5 (SLC38A5) and protein kinase binding genes stratifin (SFN)[518]. Genes involved in bone development and formation, such as parathyroid hormone like hormone (PTHLH) and Dickkopf Wnt signaling pathway inhibitor 3 (DKK3), were also some of the top downregulated genes[518].

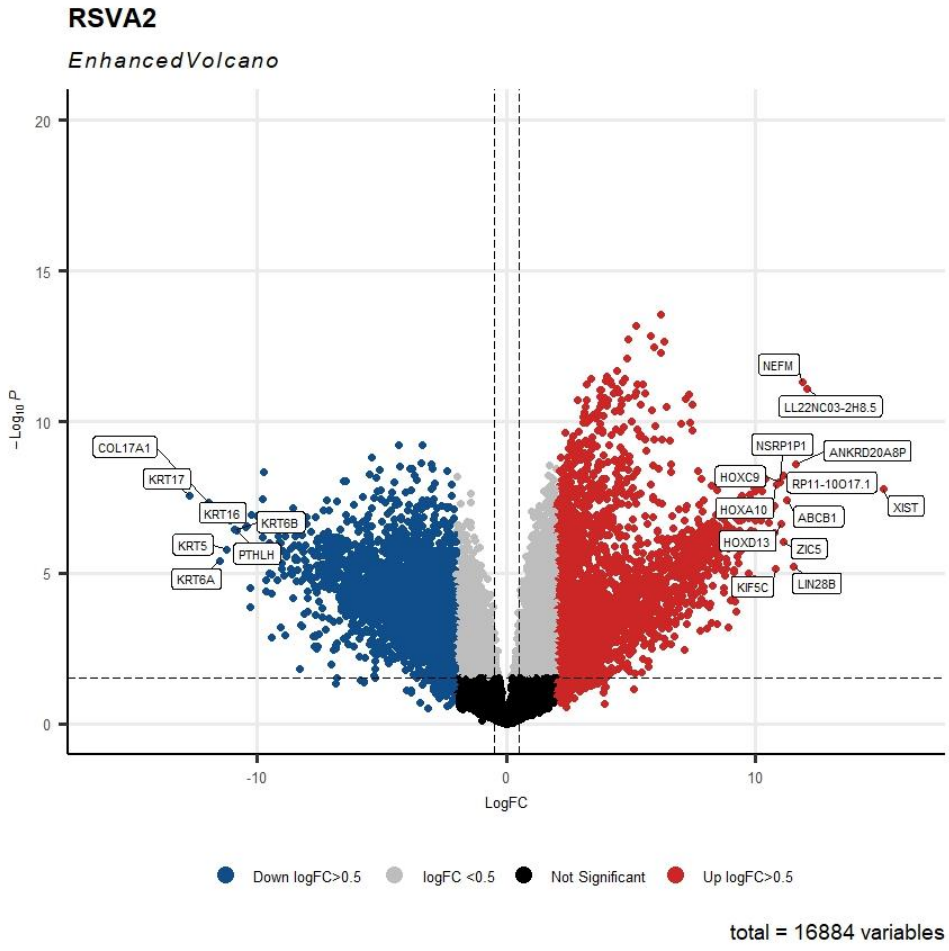


Figure 4-7. Volcano plot of differentially expressed genes in response to RSV-A2 infection after 6 days. Top 20 most significant genes are labelled. Significantly upregulated genes with logFC > 0.5 are labelled red. Significantly downregulated genes with logFC > 0.5 are labelled blue. Non-significant genes are colour-coded black and significant genes with logFC < 0.5 are colour-coded grey.

Table 4-2. Top 10 most significant upregulated genes in response to RSV-A2 infection. The table outlines the logFC and the function of the genes. Genecards and Entrez was used to obtain the function of the genes[518,704].

Gene	logFC	Function
XIST	15.0	X Inactivation
LL22NC03-2H8.5	12.0	
NEFM	11.7	Structural Molecule Activity, Structural Constituent of Cytoskeleton
LIN28B	11.6	Nucleic Acid Binding, RNA Binding
ANKRD20A8P	11.5	-
ZIC5	11.1	DNA-binding Transcription Factor Activity, RNA polymerase II-specific, Neural Crest Development
ABCB1	11.0	Transporter Activity, ATPase-coupled Transmembrane Transporter Activity

RP11-10O17.1	11.0	-
HOXD13	11.0	DNA-binding Transcription Factor Activity, Chromatin Binding
HOXC9	10.8	Sequence-specific DNA binding

Table 4-3. Top 10 most significant downregulated genes in response to RSV-A2 infection. The function of the genes and the logFC is summarised on the table. Gene functions were obtained using Genecards and Entrez[518,704].

Gene	logFC	Function
KRT17	-12.8	Structural Molecule Activity, Protein Synthesis, Epithelial Cell Growth, Tissue Repair, MHC Class II Receptor Activity
COL17A1	-12.1	Attachment of Basal Keratinocytes to Basement Membrane
KRT6A	-11.6	Structural Molecule Activity, Structural Constituent of Cytoskeleton, Wound Healing, Regulation of Epithelial Migration
KRT5	-11.3	Structural Molecule Activity, Scaffold Protein Binding, Formation of Keratin Intermediate Filaments, Maintenance of Skin Barrier
KRT16	-11.2	Structural Molecule Activity, Structural Constituent of Cytoskeleton, Regulator of Innate Immunity in Response to Breach of Skin Barrier

KRT6B	-11.0	Structural Molecule Activity, Structural Constituent of Cytoskeleton
PTHLH	-10.9	Endochondral Bone Development Protein Binding, Epithelial-mesenchymal Interactions, Hormone Activity, Peptide Hormone Receptor Binding
SLC38A5	-10.7	Amino Acid Transmembrane Transporter Activity, Glycine Transmembrane Transporter Activity, Cotransporter of Neutral Amino Acids and Sodium Ions
DKK3	-10.6	Vertebrate development, Bone Formation
SFN	-10.5	Protein Kinase Binding, Phosphoprotein Binding

4.4.1.3 Differential Expression of Genes in Response to Prolonged RV-A16 Infection

Infection with RV-A16 resulted in 4527 differentially expressed genes. 2398 genes were upregulated, and 2129 genes were downregulated. Volcano plot shows the differentially expressed genes with the top 20 most significant genes for RV-A16 with a fold change of 2 being labelled (Figure 4-8). Genes higher up on the plot have lower p-values. Statistically significant genes are colour-coded black if they did not reach logFC > 0.5. Genes that were non-significant genes were colour-coded grey. Similar to RSV-A2, XIST was the topmost significant gene. Protein homodimerisation genes such as cerberus 1 (CER1), glutathione S-transferase Mu 1 (GSTM1), and cell adhesion molecule 1 (CADM1) were also among the top 20 most significant genes[518]. CER1 has roles in bone morphogenetic protein binding whereas GSTM1 is involved in glutathione transferase activity[518]. Functions of CADM1 include cell adhesion, cell recognition, cytokine production, and formation of synapses etc. Genes with diverse roles occupy the list of topmost upregulated genes[518]. These include cell cycle regulation gene HORMA domain containing 1 (HORMAD1), cytokine activity gene left-right determination factor 2 (LEFTY2), protein crosslinking gene transglutaminase 3 (TGM3), choline transport gene solute carrier family 44 member 5 (SLC44A5), metabolic genes cytochrome P450 family 26 subfamily A member 1 (CYP26A1)[518]. In comparison to RSV-A2 infection, a keratin-encoding gene (KRT79) is one of the most significantly upregulated genes[518]. This is interesting to note as keratin-encoding genes dominated the most significantly downregulated genes for RSV-A2.

RNA-binding genes, such as ribosomal protein S4 Y-Linked 1 (RPS4Y1), DEAD-box helicase 3 Y-linked (DDX3Y), and eukaryotic translation initiation factor 1A Y-linked (EIF1AY), are among the top downregulated genes[518]. Neuroligin 4 Y-linked (NLGN4Y), which is involved in cell adhesion and synapse formation, was also one of the most significantly downregulated genes[518]. Other genes which are among the most significantly downregulated genes include taxilin gamma Y-linked pseudogene (TXLNG2P), deubiquitinase activity gene ubiquitin specific peptidase 9 Y-linked (USP9Y), histone

demethylation gene ubiquitously transcribed tetratricopeptide repeat containing, Y-linked (UTY), serine/threonine kinase activity gene protein kinase Y-Linked pseudogene (PRKY), and transcriptional activity gene zinc finger Protein Y-linked (ZFY)[518]. The list of top 10 upregulated genes by logFC for RV-A16 is shown on table 4-4 and the top 10 downregulated genes by logFC on table 4-5.

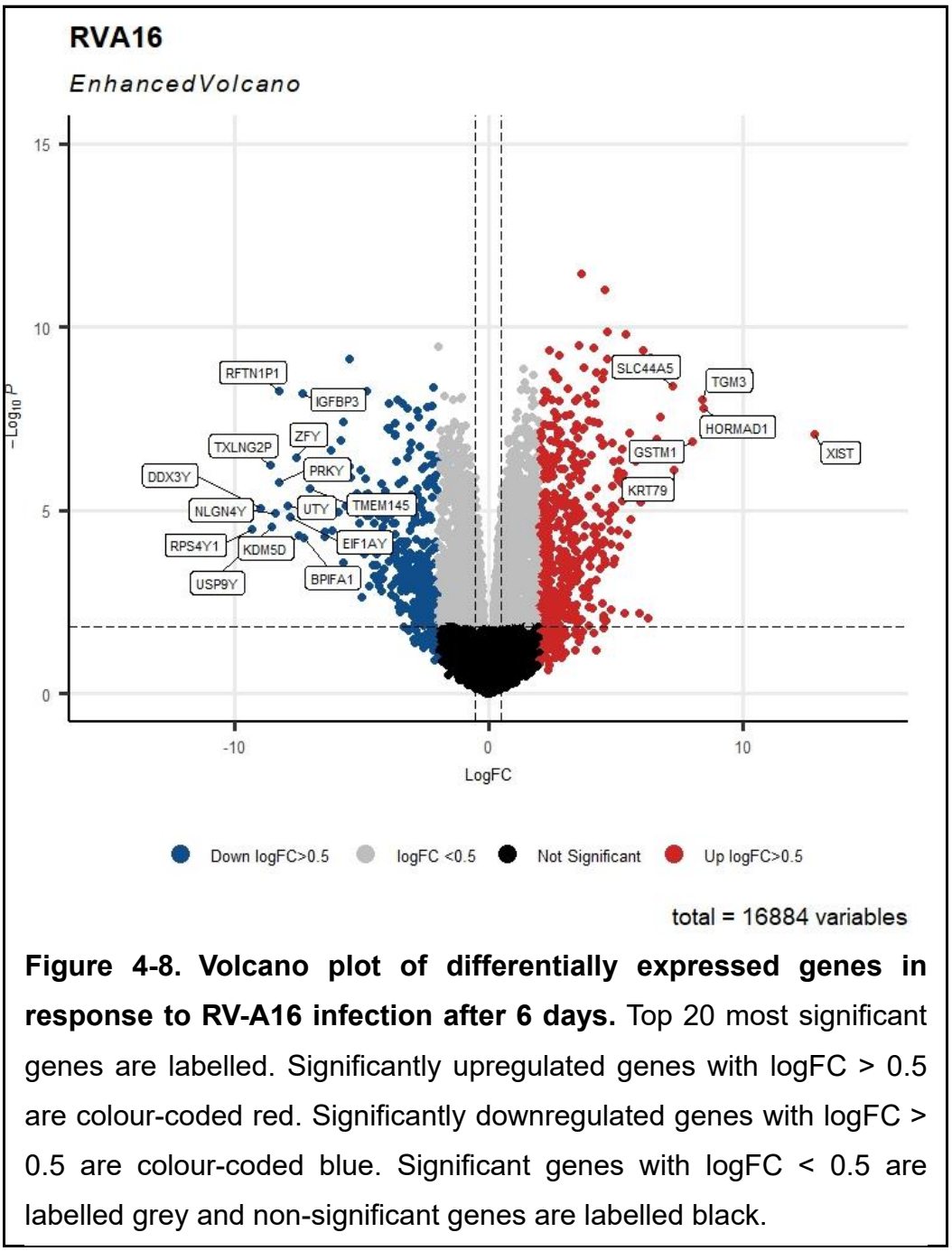


Figure 4-8. Volcano plot of differentially expressed genes in response to RV-A16 infection after 6 days. Top 20 most significant genes are labelled. Significantly upregulated genes with logFC > 0.5 are colour-coded red. Significantly downregulated genes with logFC > 0.5 are colour-coded blue. Significant genes with logFC < 0.5 are labelled grey and non-significant genes are labelled black.

Table 4-4. Top 10 most significant upregulated genes in response to RV-A16 infection. The logFC and the function of the genes are outlined on the table. Genecards and Entrez was used to obtain the function of the genes[518,704].

Gene	logFC	Function
XIST	12.8	X Inactivation
HORMAD1	8.5	Cell Cycle Regulation, Chromatin Binding
CER1	8.4	Protein Homodimerisation, Bone Morphogenetic Protein Binding
TGM3	8.4	Calcium Ion Binding, Protein-glutamine Gamma-Glutamyltransferase Activity, Calcium-dependent Formation of Isopeptide
GSTM1	8.0	Protein Homodimerisation Activity, Glutathione Transferase Activity, Formation of Novel Hepoxilin Regioisomers
LEFTY2	7.7	Cytokine Activity, Transforming Growth Factor Beta Receptor Binding

SLC44A5	7.2	Choline Transmembrane Transporter Activity, Choline Transport
KRT79	7.2	DNA-binding Transcription Factor Activity, Obsolete Signal Transducer Activity
CADM1	6.7	Signalling Receptor Binding Activity, Cell Recognition, Cytokine Production, Protein Homodimerization Activity, PDZ Domain Binding, Formation of Dendritic Spines, Synapse Assembly, Spermatogenesis
CYP26A1	6.6	Iron Ion Binding, Oxidoreductase activity, Metabolism of Retinoates

Table 4-5. Top 10 most significant downregulated genes in response to RV-A16 infection. The function of the genes and the logFC is summarised on the table. Gene functions were obtained using Genecards and Entrez[518,704].

Gene	logFC	Function
RPS4Y1	-9.2	RNA Binding
DDX3Y	-8.8	ATP Binding, RNA Binding, Hydrolysis, Formation of Intramolecular Interactions
TXLNG2P	-8.5	Syntaxin Binding
USP9Y	-8.4	Deubiquitinase Activity
NLGN4Y	-8.3	Hydrolase Activity, Cell Adhesion Molecule Binding, Formation of Synapses
RFTN1P1	-8.2	
PRKY	-8.1	Protein Serine/Threonine Kinase Activity
UTY	-7.7	Protein-protein Interaction, Demethylation of Histone H3K27me2/H3K27me3, Dioxygenase Activity

EIF1AY	-7.7	RNA Binding, Translations Initiation Factor Activity
ZFY	-7.5	Transcriptional Activation

4.4.2 What Are the Similarities and Differences in the Way Human Epithelial Cells Respond To Prolonged RSV-A2 and RVA-16 Infections?

To examine the similarities and differences between prolonged RSV-A2 and RV-A16 infection, the differentially expressed genes were interrogated. 592 upregulated genes were found to be common between the two viruses (Figure 4-9). XIST, NEFM and HOXC9, which were in the top 10 most significantly upregulated genes for RSV-A2, overlapped with RSV-A2. In contrast, seven genes out of top 10 most significantly upregulated genes for RV-A16 (XIST, HORMAD1, TGM3, GSTM1, SLC44A5, KRT79, CADM1) were found to overlap with RSV-A2 infection. Among the overlapping upregulated genes were several HOX genes, such as HOXC9, HOXA10 and HOXB9, which are important in embryonic development[520]. Another gene fras1 related extracellular matrix 2 (FREM2) encodes for extracellular matrix proteins that are essential for regulating the adhesion process during embryonic development between the epidermal basement membrane and the subdermal layer[521]. Neuronal pentraxin 1 (NPTX1) has various cellular and synaptic functions such as promotion of neurite outgrowth[522,523]. Whereas slain motif family member 1 (SLAIN1), which encodes for microtubule tracking proteins, is important for regulating the dynamics of cytoplasmic microtubules[524]. Another overlapping gene, protocadherin 10 (PCDH10) has a wide variety of functions such as cell adhesion, regulation of actin assembly, formation and maintenance of neural circuits and synapses[525]. Chemokine receptor 4 (CXCR4), with functions in cell migration, cell proliferation, and haematopoiesis was also found to overlap between the two viruses[526].

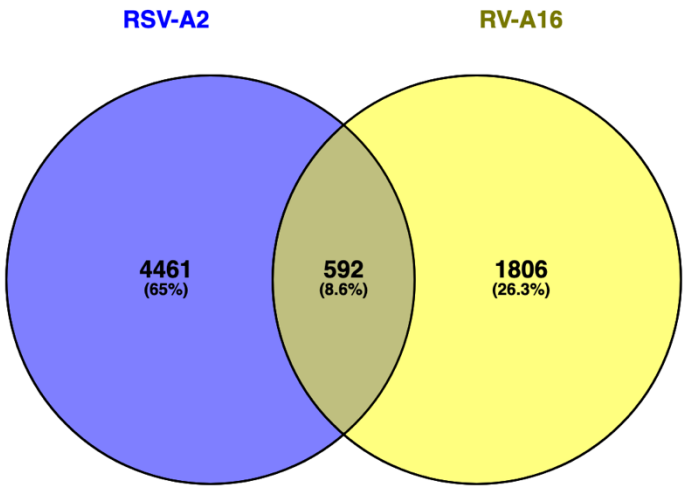


Figure 4-9. Venn diagram of overlapping upregulated genes between RSV-A2 and RV-A16. 592 upregulated genes were common between RSV-A2 and RV-A16. Venn diagram was generated using Venny[515].

Analysis of downregulated genes revealed 1125 overlapping genes between RSV-A2 and RV-A16 (Figure 4-10). Out of the top 10 most significant downregulated genes for RSV-A2, 5 were found to overlap with RV-A16: KRT17, COL17A1, SLC38A5, DKK3 and SFN. Whereas all 10 of the most significant downregulated genes for RV-A16 were found to overlap with RSV-A2. Other overlapping downregulated genes include gap junction beta-2 protein (GJB2) encodes for gap junction protein connexin 26 which is implicated in hearing loss[527]. Follistatin (FST) neutralises TGF-β superfamily such as activin, bone morphogenetic protein and myostatin to regulate muscle and fat mass[528]. Growth arrest-specific 6 (GAS6) is another overlapping downregulated gene. GAS6 maintains numerous cellular processes such as proliferation, differentiation, migration, adhesion, and apoptosis[529]. Another gene teneuricin-2 (TENM2) encodes for cell surface proteins and is known to regulate synaptic connections[530]. An important regulator of haemoglobin function, carbonic anhydrase XII (CA12) catalyses the hydration of carbon dioxide to produce carbonic acid and maintain homeostasis[531].

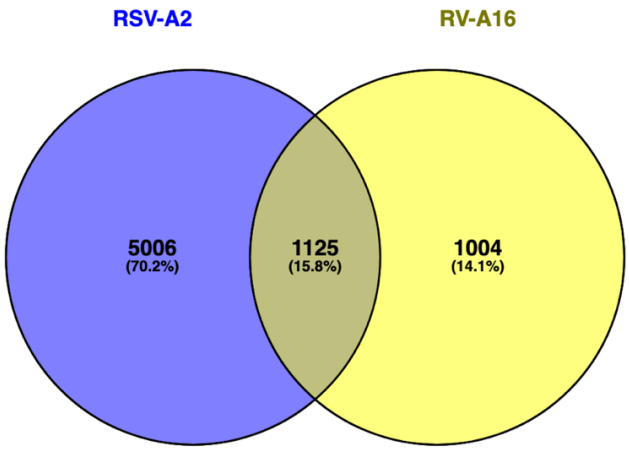
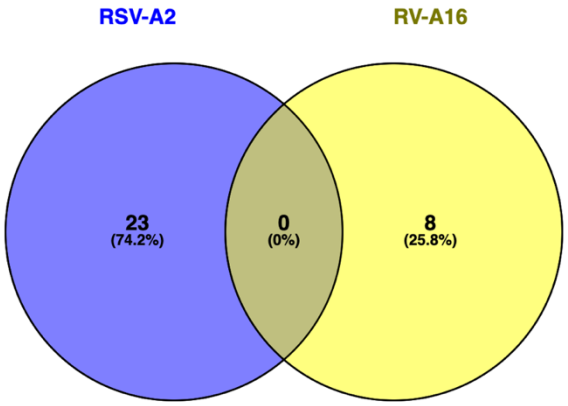


Figure 4-10. Venn diagram of overlapping downregulated genes between RSV-A2 and RV-A16. 1125 downregulated genes overlapped between RSV-A2 and RV-A16. Venn diagram was generated using Venny[515].

The differentially expressed genes were interrogated using Ingenuity Pathway Analysis (IPA). IPA predicts the directional effects of identified pathways by assigning a z-score[532]. In IPA, a z-score is an activation score and considers the known directional effects of one molecule on another molecule or process, and the change in direction of the molecules[532]. Positive z-score indicates increase in activity whereas negative z-score indicates a decrease in activity[532]. For RSV-A2 infection, IPA analysis identified 536 pathways. Although 93 pathways were assigned a positive z-score, 23 pathways out of the 93 possessed a z-score of > 2. This implies that 23 pathways were predicted to show significant activation. Cell cycle checkpoints were the top pathway with a z-score of 4.75 (Figure 4-11B). Among the pathways with z-score of > 2 were those implicated in cell cycle regulation such as cell cycle checkpoints, mitotic prometaphase, activation of the pre-replicative complex and kinetochore metaphase signaling pathway (Figure 4-11B).

Other pathways with predicated significant activation included peroxisome proliferator-activated receptors (PPAR) signalling, nucleosome assembly and regulation of TP53 activity through phosphorylation. For RV-A16 infection, downstream analysis revealed 206 pathways. 77 pathways had a positive z-score. However, only 8 out of 77 pathways had a z-score of > 2 which suggests significant activation of the pathway. Keratinisation was the top pathway with the highest z-score of 5. Pathways with z-score of > 2 included retinoic acid receptor (RAR) activation, macrophage stimulating protein (MSP)- recepteur d'Origine Nantais (RON) signalling in macrophages pathway, sperm motility, intrinsic prothrombin activation pathway, neuroprotective role of thimet oligopeptidase 1 (THOP1) in Alzheimer's disease, matrix metalloproteases inhibition, and synaptic long-term depression. However, there was no overlap in pathways with predicted significant activation of z-score > 2 between RSV-A2 and RV-A16 (Figure 4-11A).

A)



B)

RSV-A2	RV-A16
Cell Cycle Checkpoints	Keratinisation
RHOGDI Signaling	RAR Activation
Mitotic Prometaphase	MSP-RON Signaling in Macrophages Pathway
Activation of the Pre-Replicative Complex	Intrinsic Prothrombin Activation Pathway
Kinetochore Metaphase Signaling Pathway	Sperm Motility
PPAR Signaling	Neuroprotective Role of THOP1 in Alzheimer's Disease
Nucleosome Assembly	Synaptic Long-Term Depression
Cardiac β-adrenergic Signaling	Inhibition of Matrix Metalloproteases
Regulation of TP53 Activity through Phosphorylation	
HDR through Homologous Recombination (HRR) or Single Strand Annealing (SSA)	

Figure 4-11. A) Venn diagram showing the overlap in pathways with predicted significant activation between RSV-A2 and RV-A16. B) Table showing the top pathways with significant activation for RSV-A2 and RV-A16. Venn diagram created using Venny[515]. IPA analysis was conducted to identify canonical pathways and ordered based on z-score. Z-score of > 2 suggests significant activation.

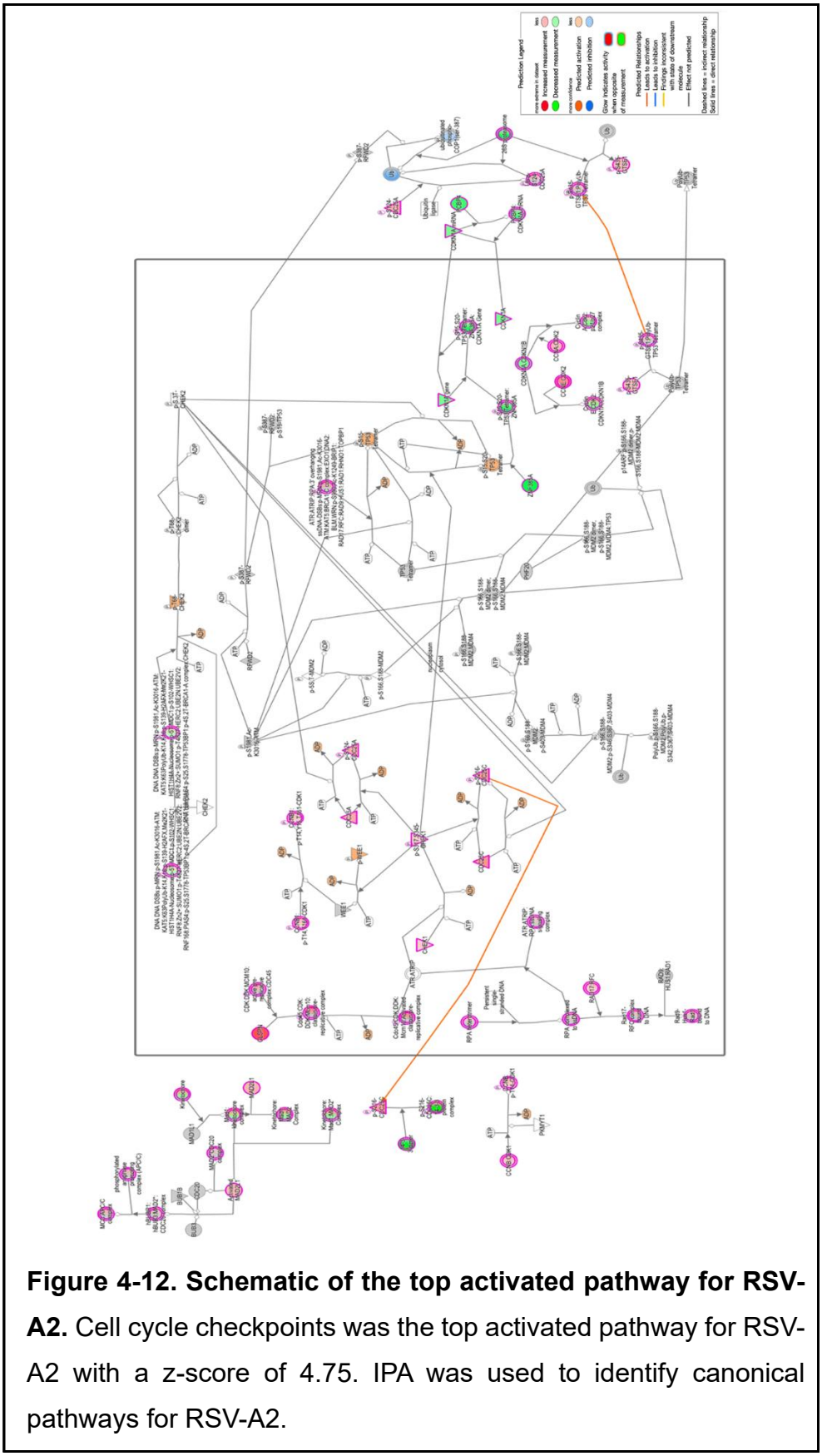


Figure 4-12. Schematic of the top activated pathway for RSV-A2. Cell cycle checkpoints was the top activated pathway for RSV-A2 with a z-score of 4.75. IPA was used to identify canonical pathways for RSV-A2.

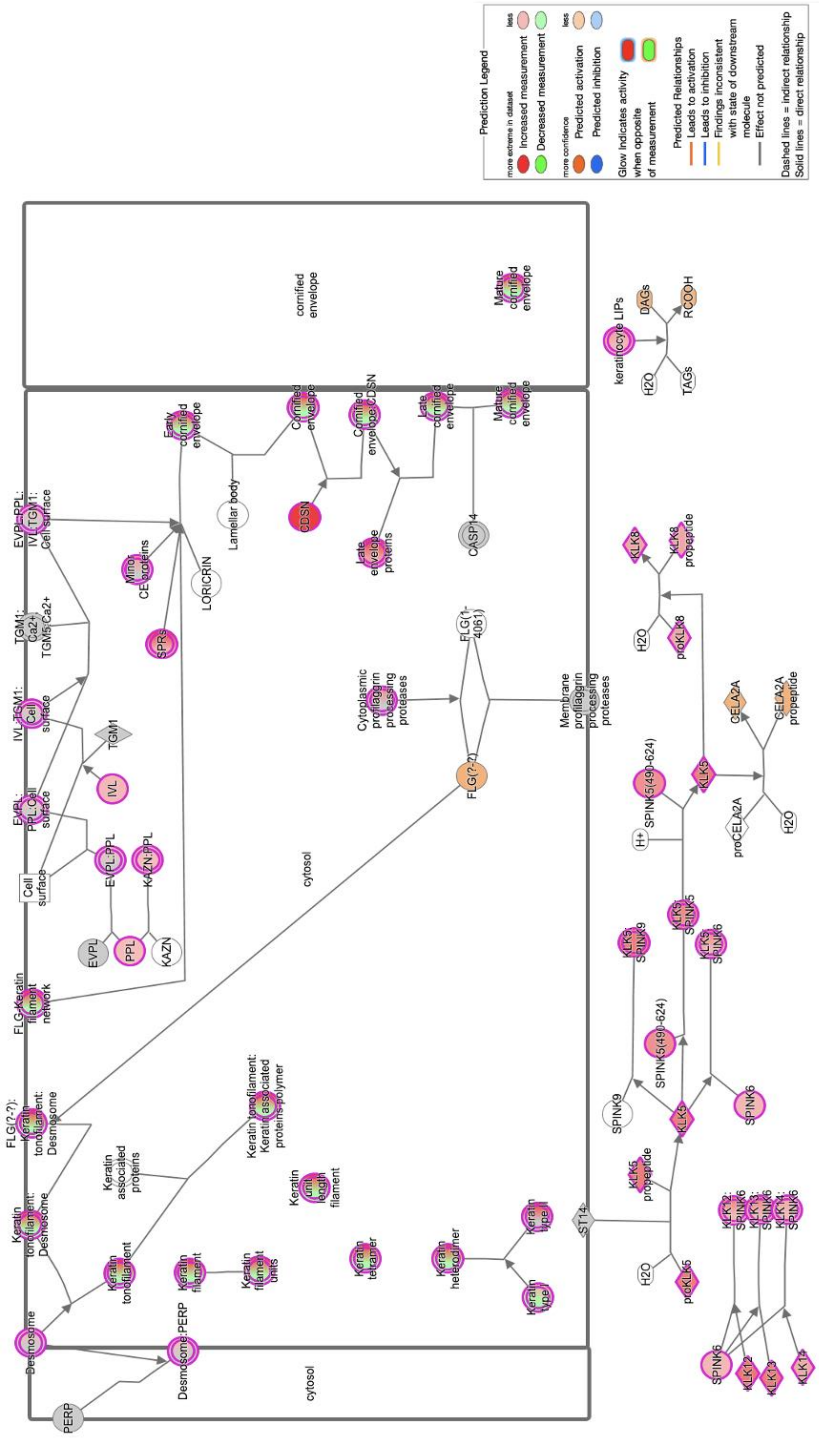
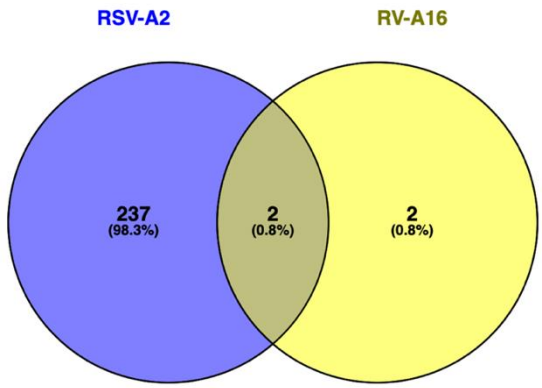


Figure 4-13. Schematic of the top activated pathway for RV-A16. Keratinisation was the top activated pathway for RSV-A2 with a z-score of 5. IPA was used to identify canonical pathways for RV-A16.

IPA analysis revealed 400 pathways with a negative z-score for RSV-A2. Out of 400, 239 pathways possessed a z-score of < -2 suggesting significant inhibition. The neutrophil degranulation pathway was the top inhibited pathway with a z-score of -7.47 (Figure 4-14B). Among pathways with predicted significant inhibition were those involved in signal transduction such as RHO GTPase cycle and pathways implicating the immune system e.g. interleukin-1 family signalling, acute phase response signalling and role of macrophages, fibroblast and endothelial cells in rheumatoid arthritis (Figure 4-14B). Other pathways with predicted significant inhibition include generic transcription pathway, keratinisation, cell junction organisation, irritable bowel syndrome signalling pathway and MSP-RON signalling in cancer cells pathway. For RV-A16, IPA analysis revealed 50 pathways with a negative z-score. However, only 4 out 50 pathways had a negative z-score of < -2 suggesting significant inhibition. Elastic fibre formation was the top inhibited pathway with a z-score of -2.45. The following pathways were also predicted to show significant inhibition: serine protease inhibitor kazal type 1 (SPINK1) pancreatic cancer pathway, potassium channels, and pulmonary fibrosis idiopathic signalling pathway. Although pulmonary fibrosis idiopathic signalling pathway and elastic fibre formation were not in the top 10 pathway with lowest z-score for RSV-A2, the two pathways were common between RSV-A2 and RV-A16 (Figure 4-14A). Whereas potassium channels and SPINK1 pancreatic cancer pathway was unique to RV-A16.

A)



B)

RSV-A2	RV-A16
Neutrophil Degranulation	Elastic Fibre Formation
RHO GTPase cycle	SPINK1 Pancreatic Cancer Pathway
Generic Transcription Pathway	Potassium Channels
Interleukin-1 Family Signaling	Pulmonary Fibrosis Idiopathic Signaling Pathway
Keratinization	
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	
Acute Phase Response Signaling	
Cell Junction Organization	
Irritable Bowel Syndrome Signaling Pathway	
MSP-RON Signaling in Cancer Cells Pathway	

Figure 4-14. A) Venn diagram showing the overlap in pathways with predicted significant inhibition between RSV-A2 and RV-A16. B) Table showing the top pathways with significant inhibition for RSV-A2 and RV-A16. Venn diagram created using Venny[515]. Canonical pathways were identified using IPA analysis and ordered based on z-score. Z-score of < -2 suggests significant inhibition.

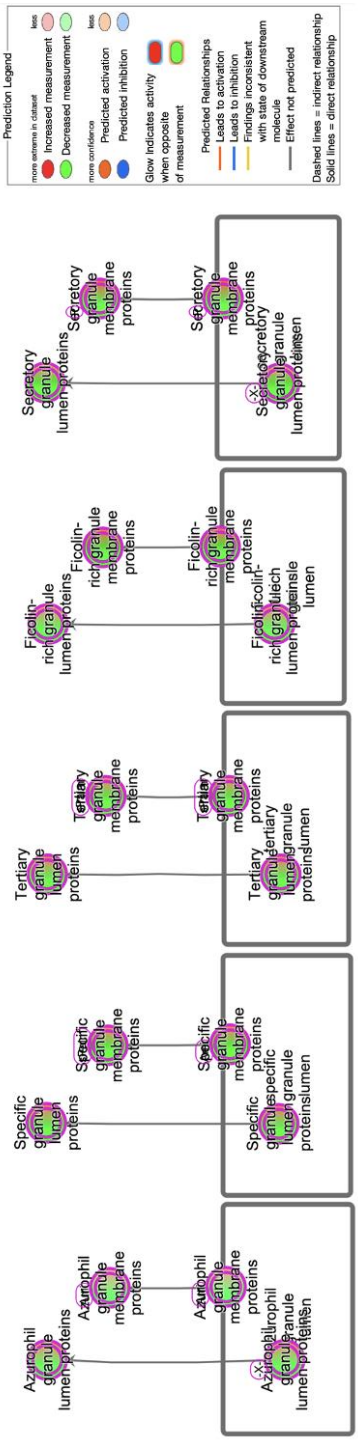


Figure 4-15. Schematic of the top inhibited pathway for RSV-A2. Neutrophil degranulation was the top inhibited pathway for RSV-A2 with a z-score of -7.47. IPA was used to identify canonical pathways for RSV-A2.

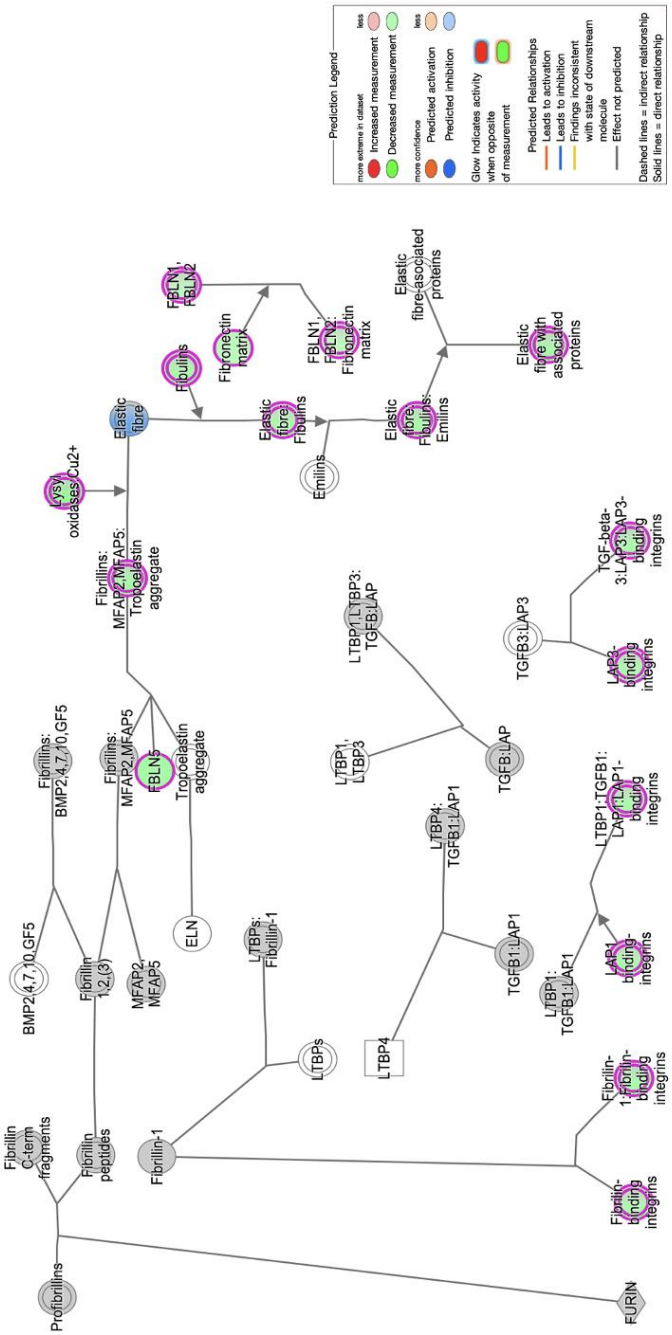


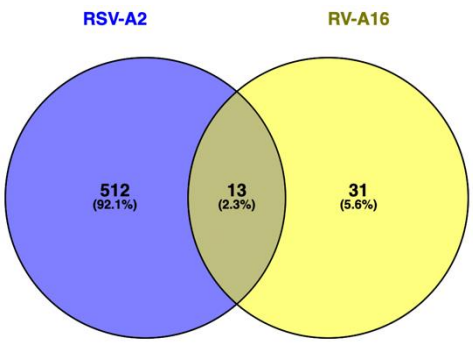
Figure 4-16. Schematic of the top inhibited pathway for RV-A16. Elastic fibre formation was the top inhibited pathway for RV-A16 with a z-score of -2.45. Canonical pathways were identified using for RV-A16.

IPA analysis was used to reveal potential upstream regulators that may drive differential change. 5372 upstream regulators were identified for RSV-A2 with 2145 possessing a positive z-score. Out of the 2145 upstream regulators, 525 were predicted to be significantly activated with a z-score of > 2. A wide range of upstream regulator types with predicted significant activation were in the top 10 for RSV-A2 including mature micro-RNA such as MiR-124-3p and MiR-3648, and enzymes like three prime repair exonuclease 1 (TREX1) and Immunity-related GTPase family M protein (IRGM) (Figure 4-17B). MiR-124-3p (and other miRNAs w/seed AAGGCAC), with a z-score of 7.05, was the top upstream regulator predicted to be significantly activated (Figure 4-17B). Other upstream regulators in the top 10 were EGFR long non-coding downstream RNA (ELDR), Ikaros family zinc finger protein 1 (IKZF1) transcription factor, protein tyrosine phosphatase receptor-type R (PTPRR) phosphatase, and the chemical drug aprepitant. Whereas IPA analysis revealed 1673 upstream regulators for RV-A16 with 441 possessing a positive-z-score. However, of the 441 upstream regulators, significant activation was predicted for only 44 molecules. The enzyme K(lysine) acetyltransferase 6A (KAT6A) was the top upstream regulator with a z-score of 3.9.

Similar to RSV-A2, a diverse range of upstream regulator types such as enzymes, fusion gene/product, chemical drug and ligand-dependent nuclear receptor, populated the top 10 for RV-A16. Furthermore, several transcription regulators such as ETS homologous factor (EHF), peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PPARGC1A), Krüppel-like factor 2 (KLF2) and transformation-related protein 63 (TP63) were predicted to show significant activation (Figure 4-17B). Other upstream regulators in the top 10 were methylprednisolone, SERPINB7, cardiotoxin and retinoic acid receptor gamma (RARG). 13 upstream regulators with predicted significant activation were common for both RSV-A2 and RV-A16. Among the overlapping upstream regulators were enzymes such as KAT6A and lipase A, lysosomal acid type (LIPA), and chemical drugs like pirfenidone and artesunic acid. Furthermore, the following transcription regulators were common for both viruses: forkhead box A1 (FOXA1), PPARGC1A, caudal type homeobox 2 (CDX2), MYOCD and MYC. Kinase inhibitor PD98059 and the fusion

gene/product EWSR1-FLI1 were also found to overlap between RSV-A2 and RV-A16. Moreover, 512 upstream regulators with predicted significant activation were unique for RSV-A2. Whereas only 31 were unique for RV-A16.

A)



B)

RSV-A2	RV-A16
MiR-124-3p (and other miRNAs w/seed AAGGCAC)	KAT6A
ELDR	EHF
RNASEH2B	Methylprednisolone
IKZF1	SERPINB7
MiR-3648 (miRNAs w/seed GCCGCGG)	NUP98-DDX10
TREX1	Cardiotoxin
PTPRR	RARG
IRGM	PPARGC1A
Aprepitant	KLF2
CKAP2L	TP63

Figure 4-17. A) Venn diagram showing the overlap in upstream regulators with predicted significant activation (z-score > 2) between RSV-A2 and RV-A16. B) Table showing the top upstream regulators with significant activation for RSV-A2 and RV-A16. Upstream regulators were colour-coded based on type: blue:mature microRNA, yellow:other, white:transcription regulator, pink:enzyme, green:phosphatase, grey = chemical drug, peach:fusion gene/product, purple:chemical- other, orange:ligand-dependent nuclear receptor.

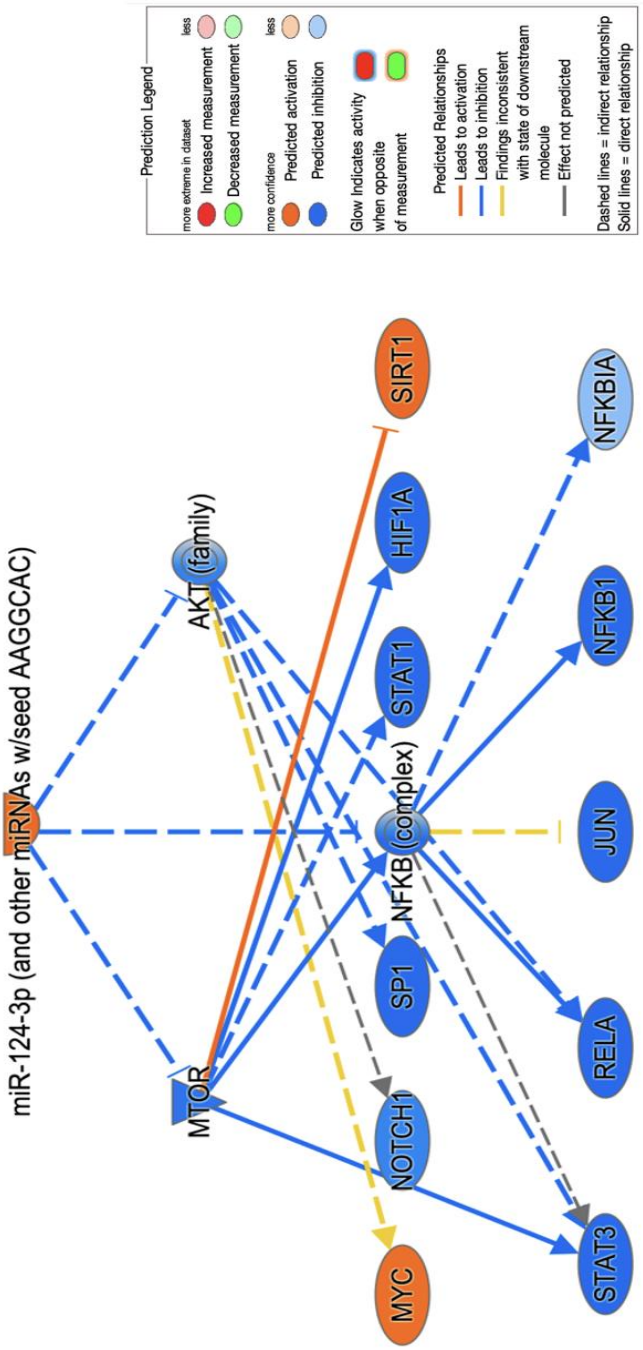


Figure 4-18. Schematic of the top activated upstream regulator for RSV-A2. miR-124-3p (and other miRNAs w/seed AAGGCAC) was the top activated upstream regulator for RSV-A2 with a z-score of 7.05. IPA was used to identify upstream regulators for RSV-A2.

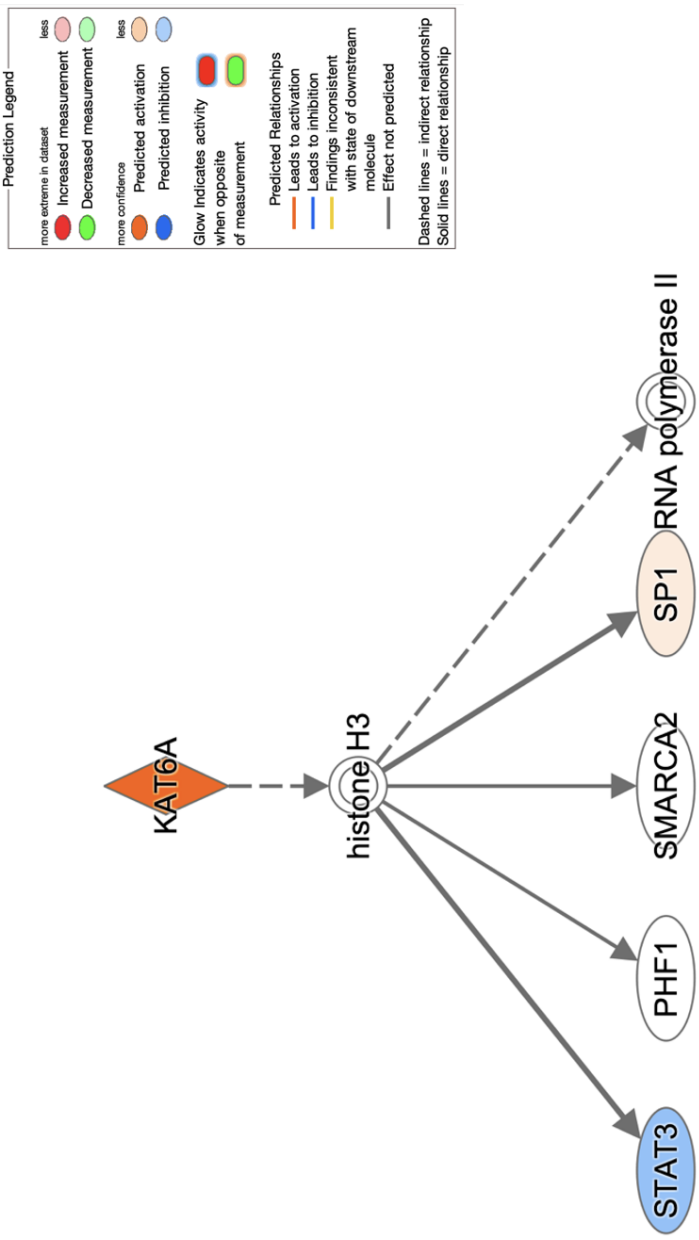
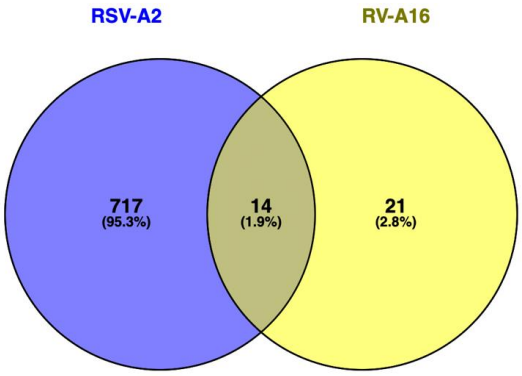


Figure 4-19. Schematic of the top activated upstream regulator for RV-A16. KAT6A was the top activated upstream regulator for RV-A16 with a z-score of 3.9. Upstream regulators was identified using IPA.

IPA analysis revealed 2333 upstream regulators possessed a negative z-score. However, 731 out of 2333 were assigned a z-score of < -2 suggesting significant inhibition. The top upstream regulator predicted to be significantly inhibited for RSV-A2 was lipopolysaccharide (LPS) (Figure 4-20B). LPS was assigned a z-score of -10.76. The top 10 included several chemical drugs (LPS, doxorubicin, tetradecanoylphorbol acetate, cisplatin) and cytokines (TNF, IL1B, IFNG) (Figure 4-20B). The transcription factor TP53 and NFKB complex were also in the top 10 (Figure 4-20B). For RV-A16, 435 upstream regulators were assigned a negative z-score. Significant inhibition was predicted for 35 out of the 435 upstream regulators since they possessed a z-score of < -2. The transcription factor SPI1 was the top upstream regulator predicted to be significantly inhibited for RV-A16 (Figure 4-20B). Furthermore, numerous transcription regulators such as myocardin-related transcription factor B (MRTFB), polycomb complex protein BMI-1 (BMI1), forkhead box C1 (FOXC1) and SRY-box 7 (SOX7) were predicted to be significantly inhibited for RV-A16.

Other upstream regulators in the top 10 were immunoglobulin G (IGG) complex, aldosterone, IKBKG kinase, fusion gene/product NUP98-KMT2A, and the chemical drug doxorubicin (Figure 4-20B). 14 upstream regulators with predicted significant inhibition were found to overlap between RSV-A2 and RV-A16. This included the kinases inhibitor of nuclear factor kappa B kinase regulatory subunit gamma (IKBKG) and mitogen-activated protein kinase kinase kinase 8 (MAP3K8), and the transcription regulators MRTFB, SPI1 and Iroquois homeobox 3 (IRX3). The overlapping upstream regulator also included the enzyme PRKCSH beta subunit of glucosidase II (PRKCSH), the cytokine TNF superfamily member 15 (TNFSF15) and the fusion gene/product NUP98-KMT2A. Chemical regulators such as aldosterone, doxorubicin, and thioacetamide were also common between the two viruses. In addition, 717 upstream regulators predicted to show significant inhibition were found to be unique for RSV-A2. In contrast, 21 upstream regulators were identified to be unique for RV-A16.

A)



B)

RSV-A2	RV-A16
Lipopolysaccharide	SPI1
TNF	MRTFB
TP53	IGG (complex)
Doxorubicin	Aldosterone
IL1B	BMI1
Tetradecanoylphorbol Acetate	FOXC1
IFNG	IKBKG
NFKB (complex)	NUP98-KMT2A
Cisplatin	SOX7
AGT	Doxorubicin

Figure 4-20. A) Venn diagram showing the overlap in upstream regulators with predicted significant inhibition (z-score < -2) between RSV-A2 and RV-A16. B) Table showing the top upstream regulators with significant inhibition for RSV-A2 and RV-A16. Venn diagram created using Venny[515]. Upstream regulators were identified using IPA analysis. The upstream regulators are colour-coded based on type: grey:chemical drug, purple:cytokine, white:transcription regulator, blue = complex, pink:growth factor, yellow:chemical endogenous mammalian, green:kinase, peach:fusion gene/product.

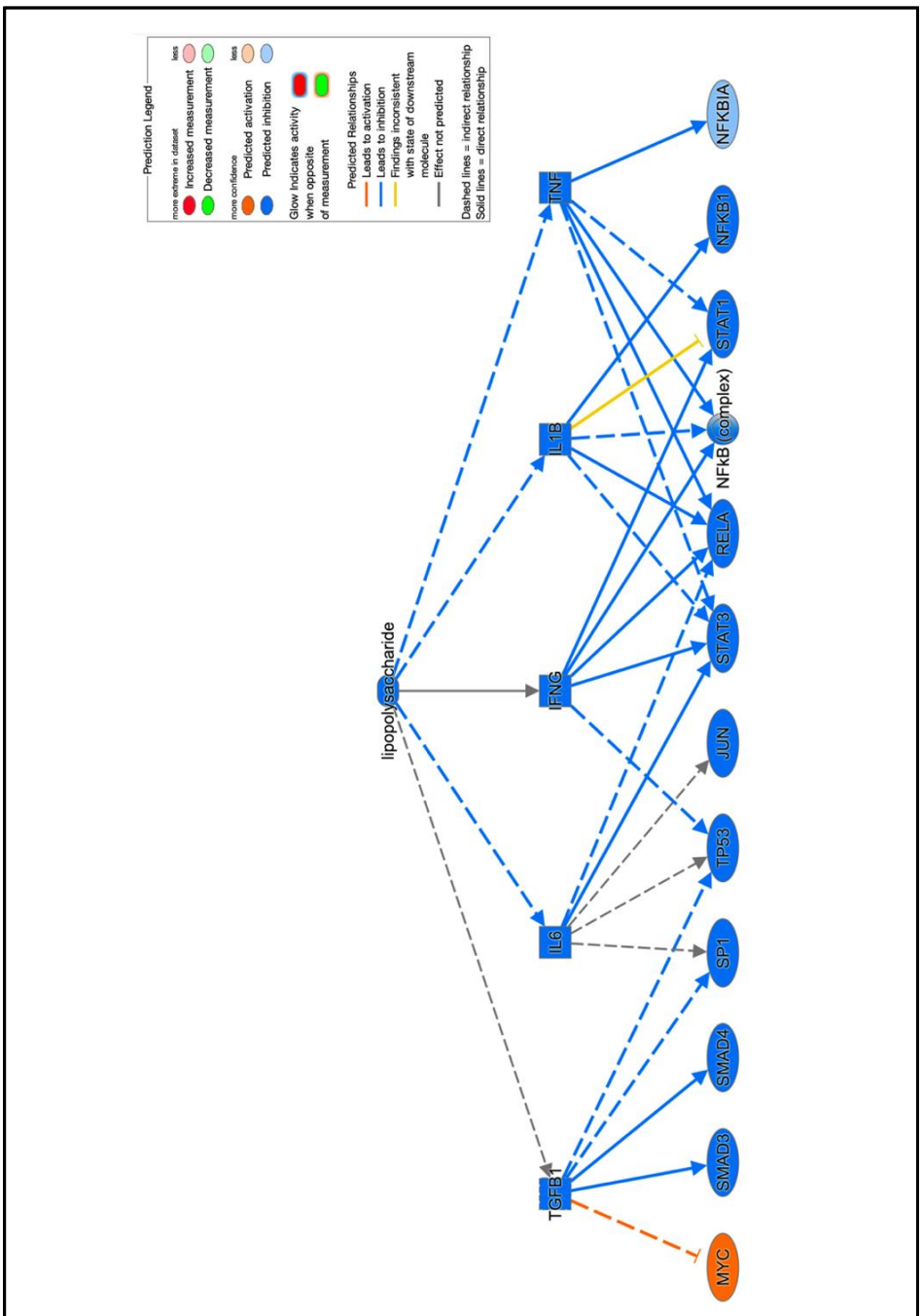


Figure 4-21. Schematic of the top inhibited upstream regulator for RSV-A2. Lipopolysaccharide was the top inhibited upstream regulator for RSV-A2 with a z-score of -10.76. IPA was used to identify upstream regulators.

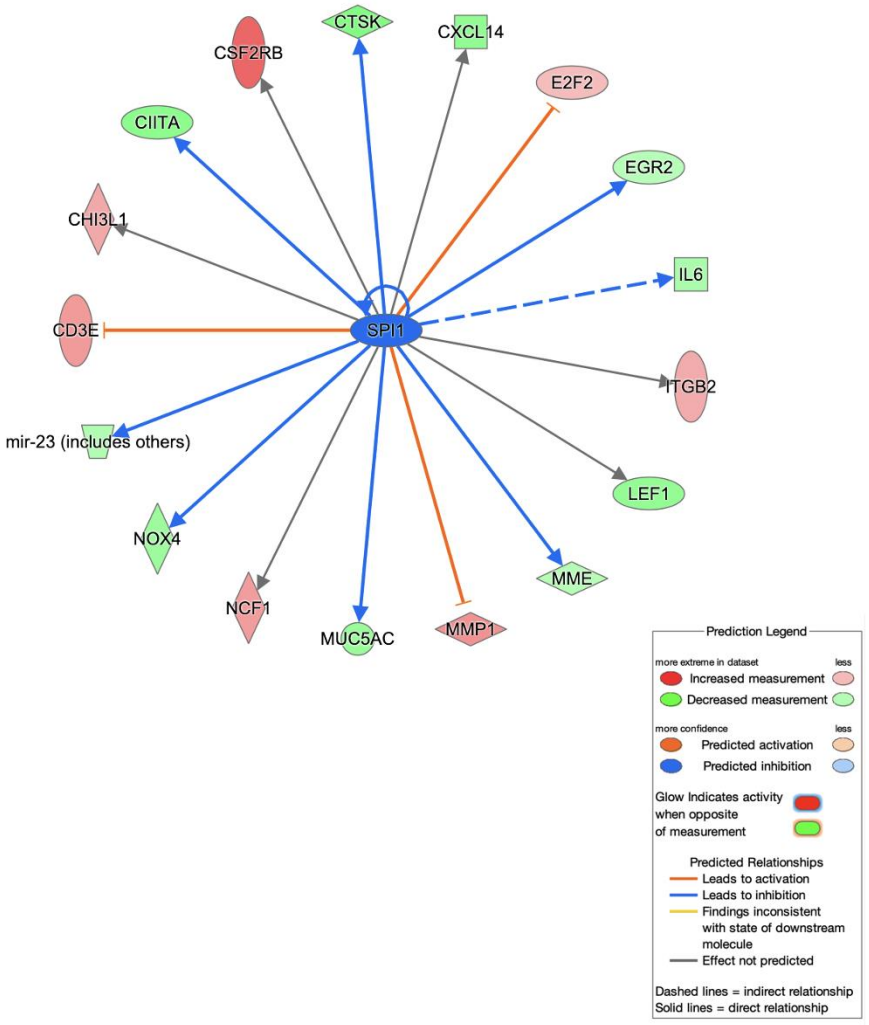


Figure 4-22. Schematic of the top inhibited upstream regulator for RV-A16. SPI1 was the top inhibited upstream regulator for RV-A16 with a z-score of -3.23. Upstream regulators were identified using IPA.

4.4.3 Does The Viral Gene Signature Provide Support For Genes Supported By Genetics?

Candidate causal genes that were identified from signal-to-gene analysis (Chapter 3) were analysed further. Published literature that conducted transcriptomic profiling of acute infection was used to assess whether candidate causal genes identified from the aforementioned analysis were differentially expressed during acute infection. A study by Scott *et al* (GEO DataSet: GSE207923) was used to obtain differentially expressed genes in response to acute 24-hour SARS-CoV-2 infection of submerged HBECs[533]. Analysis was conducted to see whether the 119 candidate causal genes associated with SARS-CoV-2 were identified in acute SARS-CoV-2 response. A caveat of this analysis is that prolonged SARS-CoV-2 infection was not conducted due to time constraints. Therefore comparison in differentially expressed genes between acute and prolonged SARS-CoV-2 infection could not be made. List of differentially expressed genes following acute RV and RSV infection in submerged HBECs were not available in published literature and so comparison between acute and prolonged infection for RV and RSV could not be made.

Out of 119 candidate causal genes identified for SARS-CoV-2 using signal-to-gene association analysis from Chapter 3, 28 genes were differentially expressed *in vitro* (Figure 4-23). This included Rho GTPase activating protein 27 (ARHGAP27) which encodes for GTP metabolising enzymes[518], ATP11A which encodes for phospholipid transporter[518] (Figure 4-23). Genes implicated with the immune response such as interferon alpha and beta receptor subunit 1 (IFNAR1), interferon alpha and beta receptor subunit 2 (IFNAR2) and OAS3 also showed association with acute SARS-CoV-2 response (Figure 4-23). Moreover, another gene NFKB inhibitor zeta (NFKBIZ), exerts an important role in various biological processes such as individual immunity and inflammatory response[534]. A number of human leukocyte antigen (HLA) genes such as HLA-A, HLA-C, HLA-DQA1, HLA-G and MICA were also found to be associated with acute SARS-CoV-2 response.

Genes with a diverse range of functions were found to be associated with acute SARS-CoV-2 infection. These included FOXP4 which is involved in lung and neuronal cell differentiation[535], FUT2 which is implicated in the secretory status of ABO antigens[536], PLSCR1 which regulates phospholipid movement[537], N-ethylmaleimide-sensitive factor (NSF) which has an essential role in intracellular vesicle transport[538], and pleckstrin homology domain containing protein family member 1 (PLEKHM1) which has shown to regulate fusion of autophagosome and lysosome during autophagy[539], DExD/H-box RNA helicase (DDX39B) which has functions in mRNA splicing and nuclear export[540], and RAVER1 which forms complexes with microfilament-associated proteins[448] (Figure 4-23). Keratinocyte-associated protein 2 (KRTCAP2), which encodes proteins implicated in N-glycosylation, was another gene associated with acute response[541] (Figure 4-23).

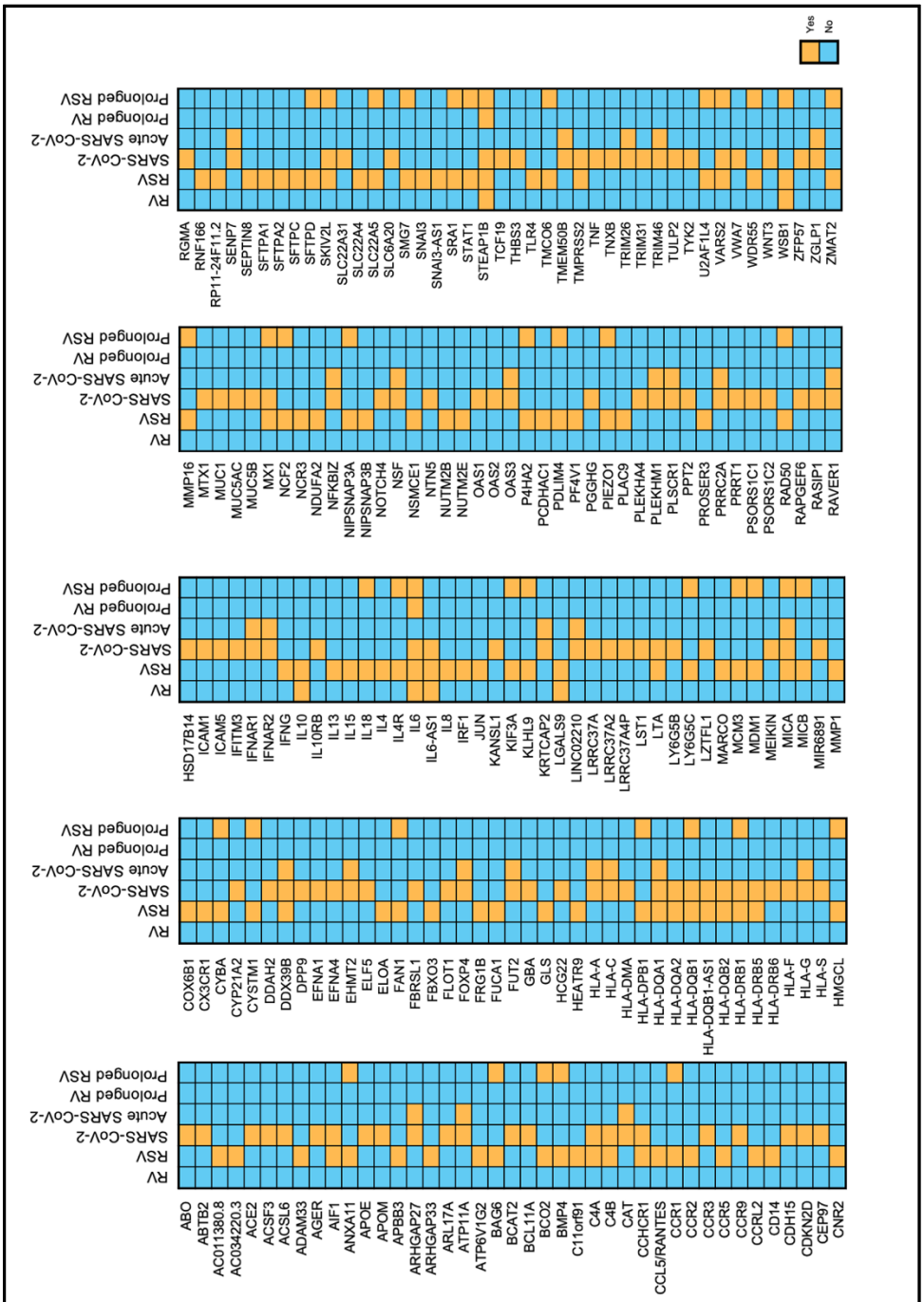


Figure 4-23. Schematic to show the candidate causal genes identified from signal-to gene association analysis in acute SARS-CoV-2, prolonged RV-A16 and prolonged RSV-A2 infection. . Orange indicates that the gene is implicated in virus susceptibility and blue are genes not associated with the virus.

Candidate causal genes that were identified from signal-to-gene analysis with 6 and 108 showing association with susceptibility to, or severity of RV and RSV disease respectively (Chapter 3) were analysed further. Out of 108 genes that were implicated in Chapter 3 as important for RSV host susceptibility, 43 were differentially expressed *in vitro* in response to prolonged RSV-A2 infection (Figure 4-24). Whereas 2 out of 6 genes identified for RV were differentially expressed following RV-A16 infection(Figure 4-24). The candidate genes were interrogated to see if their expression were either upregulated or downregulated during prolonged RV-A16 and RSV-A2 infection. The expression of IL-6 and STEAP1B were found to be significantly downregulated during prolonged RV-A16 infection (Figure 4-24). Whereas the expression of IL10, IL6-AS1, LGALS9 and WSB1 were not significant. IL-6 and STEAP1B were also significantly downregulated following prolonged RSV-A2 infection (Figure 4-24). Whereas the expression of HLA genes such as HLA-DPB1, HLA-DQB1, HLA-DRB1, and LY6G5C were significantly upregulated. Among significantly upregulated genes were those involved in carotenoid catalysis (BCO2), developmental processes (BMP4)[427], immune response (CCR1)[542], intracellular transport (KIF3A)[543], mini-chromosome maintenance (MCM3)[544], microtube binding (MDM1)[545], epithelial-to-mesenchymal transition (MMP16)[421], cell damage response (RAD50)[431], and RNA binding (U2AF1L4)[518] (Figure 4-24).

The expression of cytokine genes such as IL18 and IL4R were also found to be significantly downregulated (Figure 4-24). Interestingly genes of the HLA systems (BAG6, MICA and MICB) exhibited significant downregulation (Figure 4-24). Genes encoding Ca^{2+} dependent phospholipid-binding protein (ANXA11), oxidases (CYBA) and GTP-binding protein (MX1), ion channels (PIEZO1), surfactant protein (SFTPD), mitochondrial aminoacyl-tRNA synthetases (VARS2), and zinc finger proteins (ZMAT2) were significantly downregulated. Genes implicated in a wide range of functions were significantly downregulated such as those involved in stress tolerance (CYSTM1)[546], ketogenesis (HMGCL)[547], DNA cross-linking (FAN1)[430], superoxide production (NCF2)[548], collagen metabolism (P4HA2)[549], carnitine transport (SLC22A5)[518], non-sense mediated mRNA decay (SMG7)[435], rRNA

production (WDR55)[550], and regulation of metastatic diseases (WSB1)[416] (Figure 4-24).

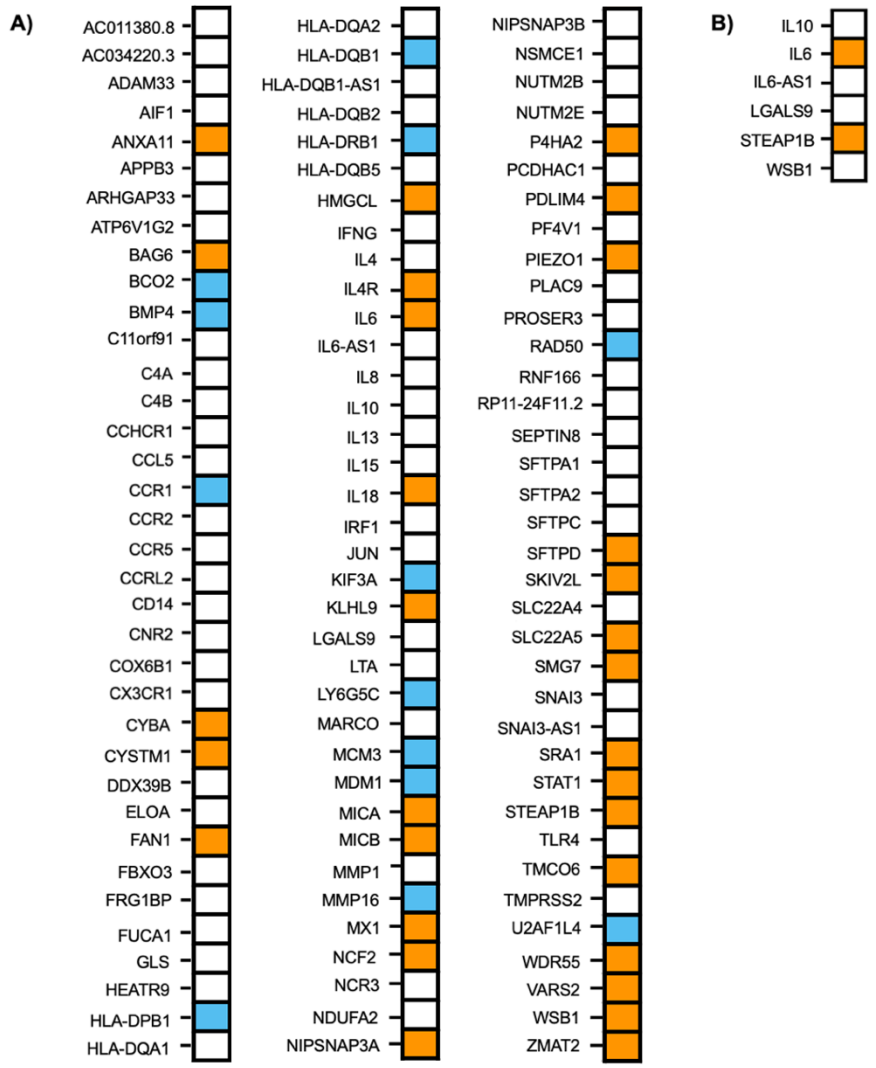


Figure 4-24. Heatmap of. A) Candidate Casual Genes Associated with RSV Infection. B) Candidate Casual Genes Associated with RV Infection. Signal-to-gene association analysis were conducted to identify candidate causal genes that showed susceptibility to or severity of RSV and RV infection as detailed on Chapter 3. The genes were examined to see if they are upregulated or downregulated during prolonged RSV-A2 and RV-A16 infection. Upregulated genes are colour-coded blue, downregulated genes are colour-coded orange, and non-significant genes are colour-coded white.

4.5 Discussion

Examining the similarities and differences in host-response to different respiratory viruses can give insights on shared host response. The aim of the Chapter was to identify and compare differentially expressed genes and pathways following prolonged RSV-A2 and RV-A16 infection. Differentially expressed genes were identified with 592 upregulated genes and 1125 downregulated genes being common for both viruses. IPA analysis revealed that pathways implicating the cell cycle may be particularly important for prolonged RSV-A2 infection. In comparison pathways involving stress and apoptosis such as keratinisation may be activated during prolonged RV-A16 infection. Significantly activated pathways showed no overlap between the two viruses which could reflect the little overlap seen between the susceptibility genes translated from signal-to-gene association analysis in Chapter 3. Pathways implicating the immune and inflammatory response such as neutrophil degranulation were predicted to be significantly inhibited during prolonged RSV-A2 infection. In addition, pathways involving elastic fibre formation and pulmonary fibrosis idiopathic were predicted to be significantly inhibited in both viruses. Interrogation of signal-to-gene association analysis revealed that the two overlapping genes IL6 and STEAP1B were found to be significantly downregulated during prolonged RSV-A2 and RV-A16 infection suggesting that they may have limited role during later stages of infection. 27 genes out of 119 identified from signal-to-gene association analysis for SARS-CoV-2 were found to be associated with acute SARS-CoV-2 response. This included genes implicated in immune response such as IFNAR1 and IFNAR2, and genes with a diverse range of roles.

XIST, a long non-coding RNA (lcRNA), was the top upregulated gene for both RSV-A2 and RV-A16. Increasing amounts of research has suggested that lncRNA may play a role in inflammation and regulation of the host immune system response[551]. In particular, XIST has been reported to modulate inflammation by regulating the Nuclear Factor-Kappa B (NF-κB) signalling pathway[551]. Several homeobox genes such as HOXD13, C9 A5, B5, B6, and B8 were found to be upregulated during prolonged RSV-A2 infection. This is in

line with literature as HOXA5, B5, B6, and B8, which play an important role in lung development, were found to be significantly upregulated following RSV-A2 infection[552]. A study by Tan *et al* found that RSV non-structural protein NS1 interacts with host histone to stimulate histone ubiquitination and subsequently induce HOX gene expression in HBECs[552]. This RSV-mediated epigenetic modification may contribute to further development of RSV-induced bronchiolitis[552].

Overlap in upregulated genes were observed between RSV-A2 and RV-A16. 592 upregulated genes were found to be common between the two viruses. One of the overlapping upregulated genes was C-X-C chemokine receptor type 4 (CXCR4). Study by Chen *et al* showed that CXCR4 along with its ligand chemokine C-X-C motif ligand 12 (CXCL12) plays an essential role in the recruitment of natural killer cells by high-mobility group box-1 (HMGB1) during later stages of RSV infection[553]. The CXCL12/CXCR4 axis is located downstream of HMGB1 and contributes to HMGB1-mediated inflammatory response, persistent airway inflammation and pathological lung damage[553]. It is important to note however that Chen *et al* used murine models and so it would be meaningful to investigate this axis using models of the human bronchial airway epithelium[553]. Another overlapping gene between the two viruses, GSTM1, has been reported to be associated with increased risk of asthma in adults and children[554]. In addition, GSTs have been connected to epithelial plasticity which is a key pathogenic characteristic of asthma[555]. Given that children who experience RV-induced wheezing at an early age are likely to develop persistent asthma, further investigation into interaction between RV and GSTs may be fruitful.

Between RSV-A2 and RV-A16, 1125 downregulated genes were found to overlap between the two viruses. Follistatin, one of the overlapping downregulated genes, acts as an inhibitor for TGF-β superfamily such as activin-A[556]. Activin-A is likely to modulate inflammatory responses and activate tissue repair programmes[557]. During asthma, activin-A signalling pathway is stimulated which induces remodelling of the airways[557]. Furthermore, activin-A is involved in the resolution of airway inflammation[557]. Considering RV infection can also promote release of activin-A from bronchial

epithelial cells, it might be interesting to explore the role of the cytokine during resolution of RV-induced airway inflammation[557]. Although some overlap in genes was observed between prolonged RSV-A2 and RV-A16, the lack of overlapping genes could be due to differences in cell tropism, molecular mechanisms, and viral strategies that the viruses employ to evade the immune system and manipulate the host.

Downstream analysis revealed 23 pathways with predicted significant activation upon prolonged RSV-A2 infection. In particular, cell cycle pathways may be important in prolonged RSV-A2 infection. Cell cycle checkpoints was the top pathway for prolonged RSV-A2 infection. These checkpoints, which are mediated by cyclins and cyclin-dependent kinases (CDKs), are surveillance mechanisms that monitor order and fidelity of events during the cell cycle[558]. Viruses are able to manipulate the cell cycle to make conditions favourable for replication and production of viral progenies[559]. A study by Wu *et al* observed that RSV infection can arrest the cell cycle in the G_0/G_1 phase to increase the production of progeny virus in HBECs after 24 hours[559]. In addition, Pastey and Lupfer reported that DNA damage or stress response was elicited by RSV 24 hours post-infection leading to the activation of cell cycle checkpoints[560]. It is interesting to note that RSV may exploit cell cycle checkpoints during both acute and prolonged infection.

Another pathway of significant activation for RSV-A2 was PPARs signalling. PPARs are ligand-activated transcription factors and have been reported to be implicated in immune and inflammatory processes[561]. Inhibition of inflammatory cell response is correlated with activation of the isoforms PPAR-α and PPAR-γ[562]. A study by Arnold *et al* reported that PPAR agonists ciglitazone, troglitazone, 15d-PGJ2 and Fmoc-Leu suppressed inflammatory suppression during RSV infection[562]. The PPAR agonists were shown to exert a protective effect through inhibition of RSV replication and downregulation of cell-mediated cytotoxicity[562]. Therefore, PPAR agonists may serve as a therapeutic intervention for treating RSV infection. Whereas 8 pathways were predicted to be significantly activated following RV-A16 infection. None of the significantly activated identified pathways overlapped with RSV-A2 infection suggesting that mechanisms unique to the two viruses are activated during

prolonged infection. This could potentially be in line with the little overlap seen between susceptibility genes in Chapter 3 signal-to-gene association analysis. Keratinisation was the top pathway with predicted significant activation. Keratins are implicated in a diverse range of pathways such as wound healing, protection from stress and apoptosis[563]. Therefore, prolonged RV-A16 infection may trigger pathways involving cytoskeletal reorganisation, stress and apoptosis.

239 pathways were predicted to be significantly inhibited upon prolonged RSV-A2 infection. Pathways involving the immune system and inflammatory response may be particularly inhibited. Neutrophil degranulation was the top pathway with significant inhibition. Neutrophils undergo receptor-mediated respiratory burst and degranulation to induce inflammation in response to chemotactic signals[564]. Interestingly PPAR signalling, where activation of PPARs is associated with reduction in inflammation, was found to be significantly activated during prolonged RSV-A2 infection. This suggests that mechanisms involved in suppression of inflammation may be activated and those associated with inflammation may be inhibited during prolonged RSV-A2 infection. Furthermore, a study by Kirsebom *et al* reported that neutrophil is recruited and activated early on during RSV infection with peak spanning 18-24 hours post-infection[565]. This suggests that neutrophil degranulation and activation may occur during acute stages of infection and become downregulated as infection progresses[565]. Although care must be taken when evaluating the results as Kirsebom *et al* used murine models of RSV infection. Another downregulated pathway was the acute phase response signalling pathway. The pathway is a rapid inflammatory response in response to infection, tissue injury or trauma[566]. Cytokines such as IL-1 and IL-6 are responsible for manifestation of acute phase response[566]. Furthermore, IL-1 family signalling, which regulates inflammatory response, was found to be downregulated. This further suggests that inflammatory pathways are inhibited during prolonged RSV-A2 infection.

Pathways involving elastic fibre formation, SPINK1 pancreatic cancer, potassium channels, and pulmonary fibrosis idiopathic signalling were reported to be significantly inhibited following RV-A16 infection. Elastic fibre

formation pathway was also predicted to be inhibited upon prolonged RSV-A2 infection. Elastic fibres maintain the structural integrity of organs including lungs against mechanical strain by providing elasticity and recoiling to tissue[567]. Defects in elastic fibre formation including degradation of elastin can contribute to a multitude of pathological conditions such as emphysema[567]. And so, prolonged RSV-A2 and RV-A16 infection may hinder formation of elastic fibres which could potentially result in the destruction of elastic fibres and lung damage. Another pathway of significant inhibition, the idiopathic pulmonary fibrosis idiopathic signalling pathway, showed overlap with RSV-A2. Idiopathic Pulmonary Fibrosis (IPF) is a rare interstitial lung disease characterised by thickening of lung tissue and fibrotic scarring[568]. Although the hypothesis that viruses may present as an etiologic agent for IPF has been suggested, the role of respiratory viruses including RSV and RV during pathogenesis is not well documented[568].

IPA analysis was conducted to identify upstream regulators that may be significantly activated during prolonged RSV-A2 infection. The mature microRNA miR-124-3p was the top upstream regulator predicted to be significantly activated for RSV-A2. miR-124-3p is involved in processes such as cell adhesion, development of the central nervous system, and tumor suppression[569]. KAT6A, the top upstream regulator with predicted significant activation for RV-A16, also overlapped with RSV-A2. Various cellular processes such as gene transcription, tissue development, and cellular senescence are regulated by KAT6A[570]. Overlap of 13 upstream regulators with predicted significant activation was observed between RSV-A2 and RV-A16. One of the overlapping upstream regulators was the Mitogen-activated protein kinase kinase (MEK1) inhibitor PD98059. MEK1 is a component of the Mitogen-Activated Protein Kinase (MAPK)- Extracellular Signal-Regulated Kinases (ERK) signalling pathway[571]. This MEK1-mediated activation of MAPK is inhibited by PD98059[572]. A study by Zaheer *et al* showed that MEK1/2-ERK pathway downregulates RV-induced C-X-C motif chemokine ligand 10 (CXCL10) production in HBECs[573]. However, application of PD98059 enhanced RV-induced CXCL10 production[573]. Proinflammatory cytokines such as CXCL10 have been known to exacerbate RV-induced airway

inflammation[573]. Whereas Tacon *et al* reported that PD98059 resulted in decreased expression of RV-induced Matrix Metalloproteinase-9 (MMP-9) protein[574]. Considering that expression of MMP-9 is increased in RV-induced asthma and the protein is implicated in airway inflammation and remodelling, it is interesting that PD98059 has different effects on CXCL10 and MMP-9 and its subsequent impact on remodelling. In contrast, Chen *et al* found that PD98059 inhibited ERK2 activity and IL-8 production during RSV infection[575].

LPS was the top upstream regulator with predicted significant inhibition for RSV-A2. A key component of the outer membrane of most Gram-negative bacteria, LPS can initiate a signalling cascade to stimulate production of chemokines and cytokines which can evoke an intense inflammatory response[576]. Although some studies have explored the relationship between acute RSV-infection and LPS, the role of TLR4 and LPS in infection remain inconclusive. Monick *et al* found that TLR4 was upregulated at 24 hours post-RSV infection[577]. This resulted in increased binding of LPS and stimulation of LPS-dependent inflammatory response[578]. In contrast, Marr *et al* reported that TLR4 activation by LPS does not play a role in RSV pathogenesis. Prolonged RSV-A2 infection may potentially lead to the inhibition of LPS to hinder inflammatory response[579]. For RV-A16, SPI1 was the top upstream regulator with predicted significant inhibition. The transcription factor has roles in processes such as myeloid cell development, complement activation, phagocytosis and microglial immune response[580]. In addition, 13 upstream regulators with predicted significant inhibition were common between RSV-A2 and RVA-16. Among them were the transcription regulators IRX3 which is vital for the function, identity and differentiation of beige adipocyte precursors[581] and PRKCSH which plays a critical role in the protein N-terminal glycosylation process[582].

Candidate causal genes that were identified from signal-to-gene association analysis were interrogated to see whether they were differentially expressed following prolonged RSV-A2 and RV-A16 infection. Genes implicated in a diverse range of functions such as immune response (CCR5), ketogenesis (HMGCL), DNA cross-linking (FAN1) and microtubule binding (MDM1) were found to be differentially expressed. The expression of two overlapping genes,

IL6 and STEAP1B, were found to be significantly downregulated for both viruses. This suggests that IL6 and STEAP1B may not play much of a role during later stages of infection. This is in line with literature as a study by Zhu *et al* reported that although bioactivity of IL-6 increased following RV infection, this activity peaked at 48-72 hours and then reduced thereafter[583]. Whereas Levitz *et al* found that induction of IL-6 by RSV is strain specific suggesting that pathogenesis and virulence may not be homogenous between different RSV strains[584]. Furthermore, Hornsleth *et al* observed that severity of RSV disease can be related to different strains[585]. Therefore, it is important to also consider factors such as strains and cell lines/models used when investigating disease pathogenesis. Little is known about the relationship between STEAP1B and the two viruses and therefore should be investigated further. Acute SARS-CoV-2 dataset was examined to assess whether the 119 candidate causal genes identified for SARS-CoV-2 were differentially expressed *in vitro* following acute response. 27 genes were found to be differentially expressed including genes involved in immune response such as IFNAR1 and IFNAR2. A study by Yaugel-Novoa *et al* reported that IFNAR1 concentration was inversely correlated with COVID-19 severity whereas IFNAR2 concentration was significantly elevated in patients with severe COVID-19[586].

The strengths of the Chapter include comparing and contrasting two important respiratory viruses RV and RSV. Studying similarities and differences between viruses can give insight into shared host response. Furthermore, this allows the identification of common and unique genes, pathways and upstream regulators which will give us greater understanding of host-virus interactions. In addition, since little literature exists on prolonged RV and RSV infection, these findings provide understanding of genes, pathways and upstream regulators that may be important during prolonged infection. One of the limitations of the Chapter is that many genes reported to be differentially expressed following prolonged infection may be related to cell homeostasis and apoptosis. This could be due to cells experiencing senescence or cell death after 6 days in culture therefore this should be taken into consideration. Another limitation was that comparison between acute and prolonged RV, RSV

and SARS-CoV-2 infection could not be made. Future studies should compare prolonged infection with acute infection to assess the similarities and differences in the genes, pathways and upstream regulators.

In summary, understanding of genes, pathways and upstream regulators associated with prolonged RSV-A2 and RV-A16 infection were provided. Furthermore, the similarities and differences between the two viruses were examined by using this transcriptomics-based approach. Targeting common genes and pathways shared by respiratory viruses can potentially lead to successful drug repurposing of therapeutics with broad anti-viral effects. Therefore these findings provide opportunities to facilitate the development of therapeutic interventions and identification of biomarkers to aid treatment strategies.

5.0 Development of Induced Pluripotent Stem Cells and Human Bronchial Epithelial Cells-Derived Models of the Airway Epithelium to Investigate Host-Virus Interactions

5.1 Introduction

As the immunological and mechanical barrier against the external environment, the airway epithelium serves as a great candidate for investigating the molecular underpinnings behind host-virus interactions[587]. Furthermore, the airway epithelium is the primary area for gaseous exchange and also exhibits high immunological activity[11]. The airway epithelium is the initial site of interaction with respiratory viruses[588]. Moreover, it is the primary site for viral replication[588]. Therefore the properties of the airway epithelium, which helps mediate the airway homeostasis and responds to pathogens including respiratory viruses, are central for limiting the recurrence of infections[588]. Using animal models to study host-virus interactions can be expensive and time-consuming[342]. They also lack physiological relevancy to the human respiratory system[587]. Animal models such as rodents do not always fully recapitulate the architecture and the physiology of the human airway epithelium[342]. Therefore they lack physiological relevancy to the human respiratory system[587]. Whereas *ex vivo* human lung samples can be difficult to obtain[587]. This has led to the use of *in vitro* cell culture for modelling the airway epithelium. The manipulation of genes and molecular pathways is one of the benefits of using cell culture models[589].

Primary epithelial cells cultured in submerged conditions have been routinely used for their low maintenance and robust capacity for proliferation *in vitro*[590]. However, cells cultured in submerged conditions are undifferentiated monolayers and therefore it remains a challenge to investigate functions such as mucociliary clearance[591]. Although respiratory viruses can infect the basal epithelial cells that lack physiological properties of the mucociliated epithelial layer, ciliated cells are the primary target for infection for viruses such as SARS-CoV-2[592]. This has resulted in the utilisation of the air-liquid interface platform to model the differentiated airway epithelium to study host-virus

interactions. By implementing the ALI technique, the barrier of the airway epithelium can be emulated[593]. However, limitations include invasive bronchoscopy technique and finite supply of cells[594]. Whereas iPSC offers unlimited capacity of expansion for iPSCs[354]. Furthermore amenable nature for genetic engineering and the generation of patient-specific signals are some of the advantages possessed by iPSCs[595]. This Chapter focuses on the development of HBEC and iPSC-derived models of the airway epithelium to investigate host-virus interactions.

5.1.1 Induced Pluripotent Stem Cells-Derived Airway Epithelium

The foundation for generating airway epithelial cells involve differentiation hiPSCs through the normal pathways of human development towards the target cell type. For airway epithelial cells this involves differentiation through definitive endoderm (DE), anterior foregut endoderm (AFE), early lung progenitors and bronchial progenitors. During early development the epiblast, the precursor of somatic lineages, is established[596]. The germ layers of the blastula are then rearranged into three gastrula layers: ectoderm, mesoderm and endoderm[597]. Both gastrulation and the development of the endoderm transpires primarily in the node-proximal region[598]. The definitive endoderm (DE) is then incorporated into a primitive tubular anlage (primitive gut tube) through morphogenetic movements[599]. Patterning of the anlage occurs to allow formation of the epithelial compartment of various organs such as lungs, liver, pancreas and the intestines[600]. The anterior foregut epithelium (AFE) originates from the regional domain of the DE[601]. The caudal region of the AFE gives rise to the lung and trachea[602]. Furthermore, AFE contributes to the formation of other organs such as the oesophagus, stomach, and salivary glands[601]. From the ventral wall of the AFE, the lung primordium arises and develops to form the trachea[603]. At early stages of lung formation, the epithelial cells remain relatively undifferentiated with simple, cuboidal-columnar cells lining the tubule[604]. The tubules then undergo branching morphogenesis where epithelial cells start to proliferate and differentiate into various cell types along the proximal-distal axis of the lung[604]. The airway epithelium lines the conducting airways and the respiratory tract[604]. The epithelial-mesenchymal

interactions contribute to lung development and branching morphogenesis[605]. This is mediated by a network of growth factors, transcription factors and their receptors[605].

Green *et al* were the first to direct differentiation of iPSC towards the lung cell fate[602]. Exposure of embryonic stem cells to high concentration of activin A (ActA) induced DE[602]. DE markers C-X-C chemokine receptor type 4 (CXCR4), c-KIT and epithelial cell adhesion molecule (EPCAM) were detected in >95% of the cells[602]. Bone morphogenic protein (BMP) and transforming growth factor beta (TGF-β) were inhibited using NOGGIN and SB-431542 respectively to generate SOX2 and forkhead box protein A2 (FOXA2) enriched cells[602]. Addition of WNT3A, keratinocyte growth factor (KGF), fibroblast growth factor protein 10 (FGF10), bone morphogenetic protein 4 (BMP4) and epidermal growth factor (EGF) induced expression of ventral markers NK2 homeobox 1 (NKX2.1), NKX2.5 , and paired box 1 (PAX1) signifying the acquisition of lung fate[602]. Whereas Longmire *et al* developed a serum-free culture to derive pure populations of NKX2.1-expressing lung progenitors in 2012[606]. Through supplementation with activin, 80% of the DE was induced[606]. Exposure to high levels of fibroblast growth factor protein 2 (FGF2) led to establishment of NKX2.1 expressing AFE cells[606]. Treatment with cyclic AMP, dexamethasone, 3-isobutyl-1-methylxanthine (IBMX) and KGF on day 22 for three days resulted in increased expression of surfactant proteins and club cells[606].

To drive specification towards lung endoderm, Mou *et al* cultured cells in a proximal induction medium consisting of B27, bone morphogenetic protein 7 (BMP7), fibroblast growth factor protein 7 (FGF7), WNT antagonist IWR-1 and Noggin[607]. 10% of the NKX2.1$^+$ cells were NKX2.1$^+$ SOX2$^+$ airway progenitor cells compared to 1-2% before treatment with media[607]. Immunodeficient mice were then injected with the cells[607]. Epithelial spheres composed of SOX2, FOXJ1, MUC5AC, club cell protein CC10 expressing NKX2.1$^+$ cells were identified[607]. Moreover lung airway progenitors from cystic fibrosis-specific human iPSCs were generated using modified RNAs (RiPSCs)[607]. Treatment with Activin and PI3K kinase inhibitor resulted in 85-90% of the cells co-expressing FOXA2 and SRY-box 17 (SOX17)[607]. Although subcutaneous engraftment of the RiPSC-derived NKX2.1$^+$ cells led to formations of spheres,

further differentiation into ciliated, club and goblet cells were not observed[607]. In 2013, Huang *et al* refined the AFE induction approach to improve the efficiency of lung field specification[608]. Following DE induction, AFE was induced in the presence of DSM/SB431542 followed by SB431542/IWP2[608]. Next, BMP, TGF-β and WNT were sequentially inhibited, and cells were cultured with CHIR99021, FGF10, KGF, BMP4 and retinoic acid (RA) until day 15[608]. Addition of dexamethasone, 8-bromo-cAMP and isobutylmethylxanthine (DCI) led to increase in surfactant protein expression[608].

Differentiation of iPSC into a functioning CFTR-expressing airway epithelium consisting of multiciliated, basal, club, goblet cells was reported by Firth *et al*[609]. Similar to Green *et al* , AFE was generated through dual inhibition of BMP and TGF-β signalling[609]. At day 17, the airway epithelial progenitors were cultured in ALI[609]. Tight junction proteins epithelial cadherin (E-Cadherin), epithelial cell adhesion molecule (EPCAM) and zona occludens-1 (ZO-1) were expressed and the expression of FOXJ1 was highly upregulated[609]. Multiple cilia bodies and projections were observed[609]. Functional expression of CFTR was also reported[609]. Wong *et al* developed a protocol for the generation of CFTR-expressing airway epithelia from human embryonic stem cells[610]. Following DE generation using activin and WNT3A, high concentration of FGF2 and sonic hedgehog (SHH) was used to promote differentiation towards AFE[610]. The cultures were then treated with FGF7, FGF10, fibroblast growth factor protein 18 (FGF18) and a low concentration of BMP4 to drive differentiation towards proximal lung cell fate[610]. To stimulate the outgrowth of lung bud, retinoic acid was added to upregulate the expression of FGF10 and then grown in ALI for further maturation[610]. CFTR was observed to be located on the apical side of the plasma membrane along with cilia and MUC5AC[610,611].

In 2017, McCauley *et al* developed a low-Wnt protocol for the generation and maturation of proximal airway organoids from NKFX-1GFP$^+$ lung progenitors[611]. Culturing the NKX2.1$^+$ progenitor cells in CHIR99021, FGF10 and KGF drove differentiation towards the lung lineage[611]. Cyclical modulation of the Wnt signalling pathway facilitated differentiation of the NKX2.1$^+$ progenitor cells into proximal airway organoids[611]. The NKX2.1$^+$ progenitors possess high levels of Wnt activation but rapidly pattern into proximal airway

lineage in low-Wnt conditions[611]. Subsequently culturing the cells in 2D ALI or organoids in Notch inhibitor DAPT resulted in the generation of multiciliated epithelial cells[611]. In contrast, Ghaedi *et al* developed a protocol where AFE was generated by sequentially exposing the DE cells to dorsomorphin/SB431542 and IWP2/SB431542 from day 5-6 and 6-7 respectively[612]. To drive differentiation into proximal airway progenitor cells, cells were exposed to BMP7, KGF, high-concentration RA, and WNT antagonist on day 15[612]. Expression of airway progenitor markers such as P63, cytokeratin 14 (CK14) and 5 (CK5) was observed. iPSCs were also cultured in conducting airways of lung tissue scaffolds[612]. Whereas an epithelial-only airway organoid protocol was developed by McCauley, Hawkins and Kotton in 2018[612]. Embryonic or iPSCs are differentiated into DE. Dual-SMAD inhibition is then initiated to pattern the PSCs into AFE[612]. To drive differentiation into NKX2.1^{+} lung progenitor fate, Wnt, BMP and retinoic acid signalling is activated[612]. The lung progenitors are expanded in a low-Wnt, FGF-containing media to specify airway epithelial progenitors[612]. Cells are resuspended briefly in 3D Matrigel matrix and replated in tissue culture dishes[612]. FGF signalling is activated by culturing the cells in media supplemented with FGF2, FGF10, cAMP, and corticosteroids[612]. This enables formation of epithelial spheres which contain differentiated airway cell types[612].

Table 5-1. Table of Lung iPSC Differentiation Protocol. The table outlines the protocol, its duration, the year of publication and the growth factors used.

Protocol	Year	Duration of Protocol	Growth Factors
Green *et al*	2011	19 days	Y27362, bFGF, ActA, BMP4, FGF, NOGGIN, SB-431542, Wnt3a, FGF10, FGF7, BMP4, EGF, ARTA
Longmire *et al*	2012	25 days	ActA, SB-431542, Wnt3a, FGF10, FGF7, BMP4, EGF, FGF2, DC, Insulin, Transferrin, Sodium Selenite
Mou *et al*	2012	12+ days	ActA, LY294002, A-83-01, FGF2, BMP4, GSK3iXV
Huang *et al*	2013	48 days	BMP4, Wnt3a, Y27632, ActA, bFGF, DSM, SB-431542, IWP2, CHIR, FGF10, KGF, Retinoic acid, DCI

Protocol	Year	Duration of Protocol	Growth Factors
Firth *et al*	2014	45+ days	ActA, Wnt3a, NOGGIN, SB-431542, BMP4, FGF2, FGF10, FGF7,
Wong *et al*	2015	40+ days	ActA, Wnt3a, FGF2, SHH, FGF7, FGF10, BMP4, FGF18
McCauley *et al*	2017	20+ days	CHIR99021, FGF10, KGF, BMP4, RA
Ghaedi *et al*	2017	40 days	ActA, Dorsomorphin, SB-431542, IWP2, RA, bFGF, BMP4, Wnt3a, KGF, BMP7, IWR-1, EGF
McCauley, Hawkins, Kotton	2018	+22 days	Dorsomorphin, SB-431542, CHIR, BMP4, Retinoic acid, FGF2, FGF10, Y-27632, Dexamethasone, cAMP/IBMX

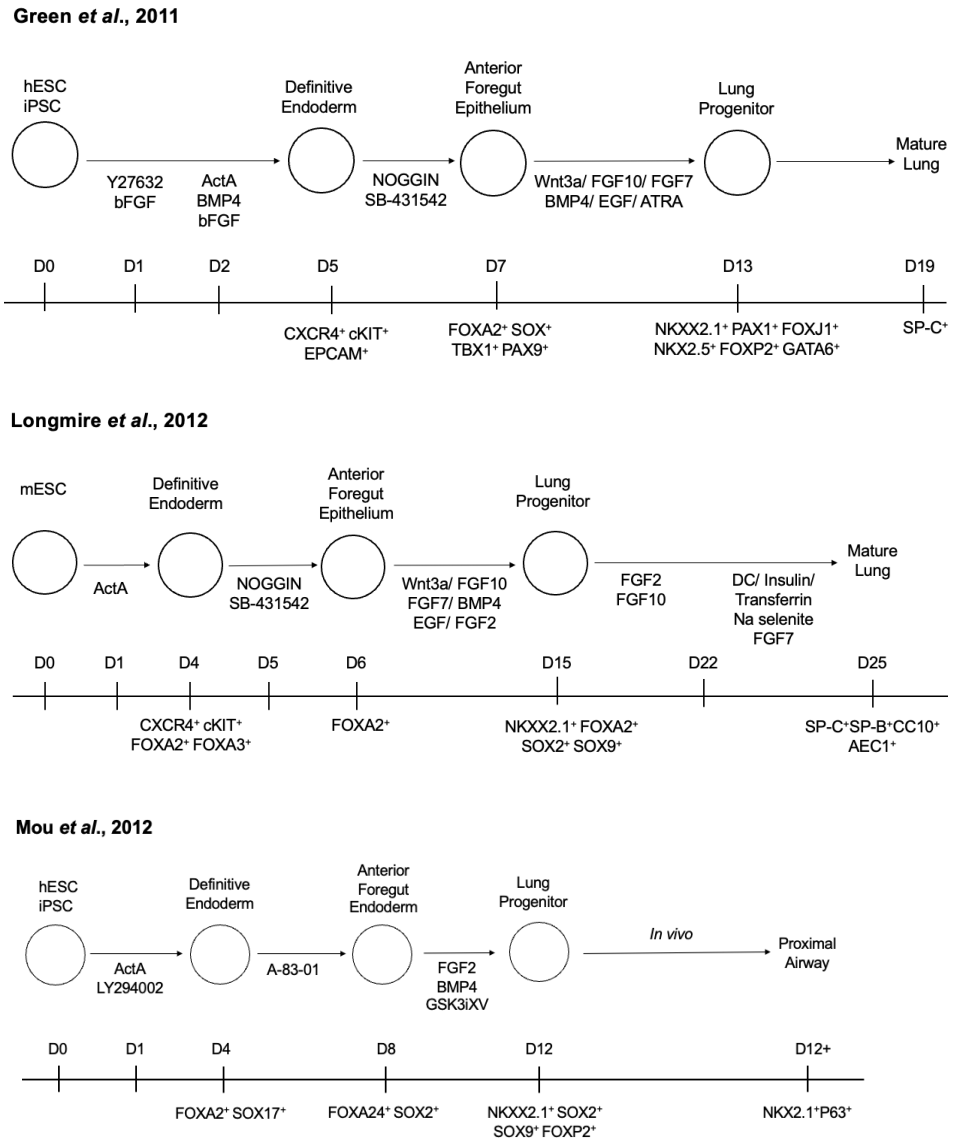


Figure 5-1. Schematic of Published Protocols on Differentiation of iPSC. Protocols by Green *et al*, Longmire *et al*, and Mou *et al* are outlined.

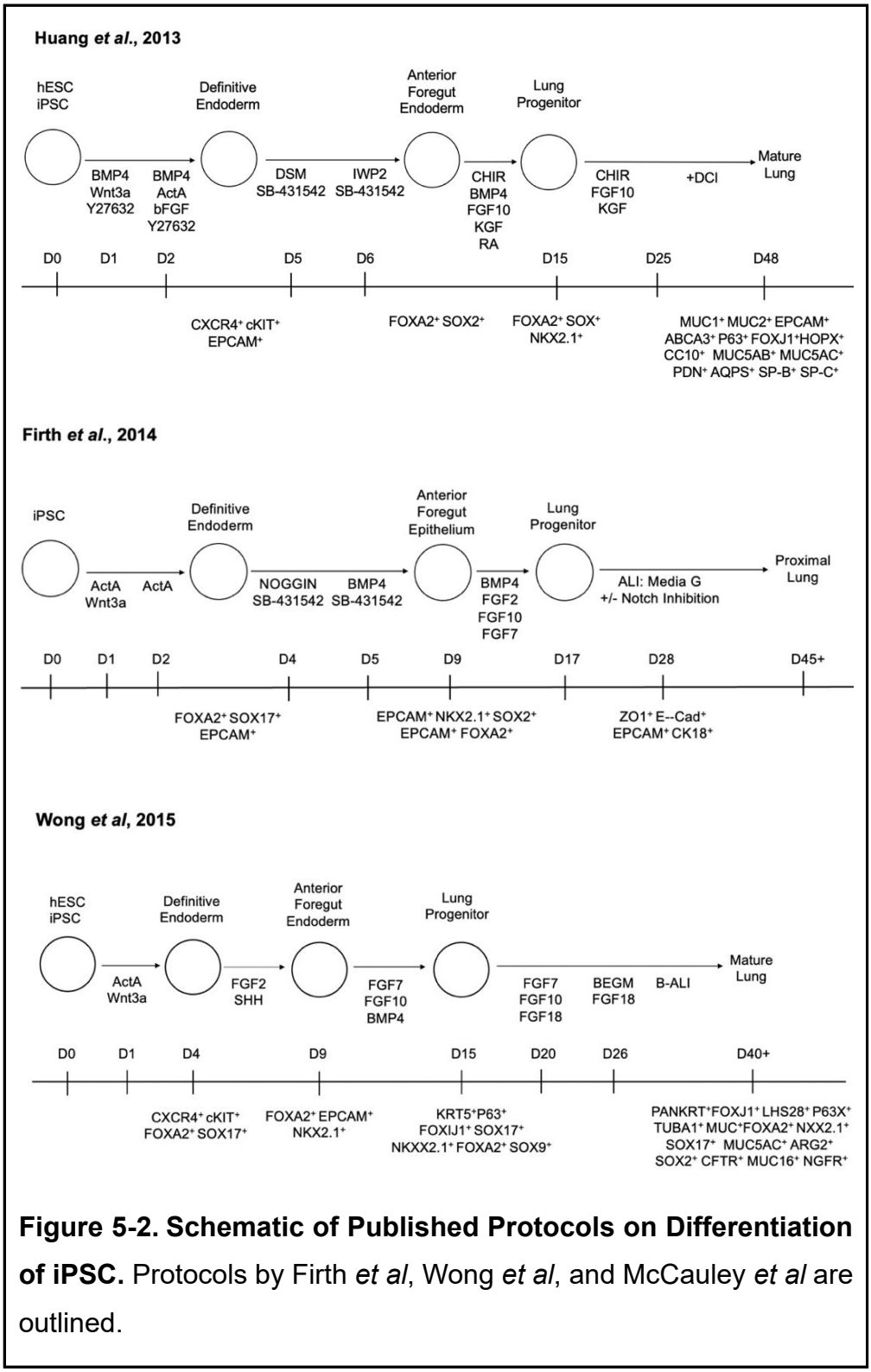


Figure 5-2. Schematic of Published Protocols on Differentiation of iPSC. Protocols by Firth *et al*, Wong *et al*, and McCauley *et al* are outlined.

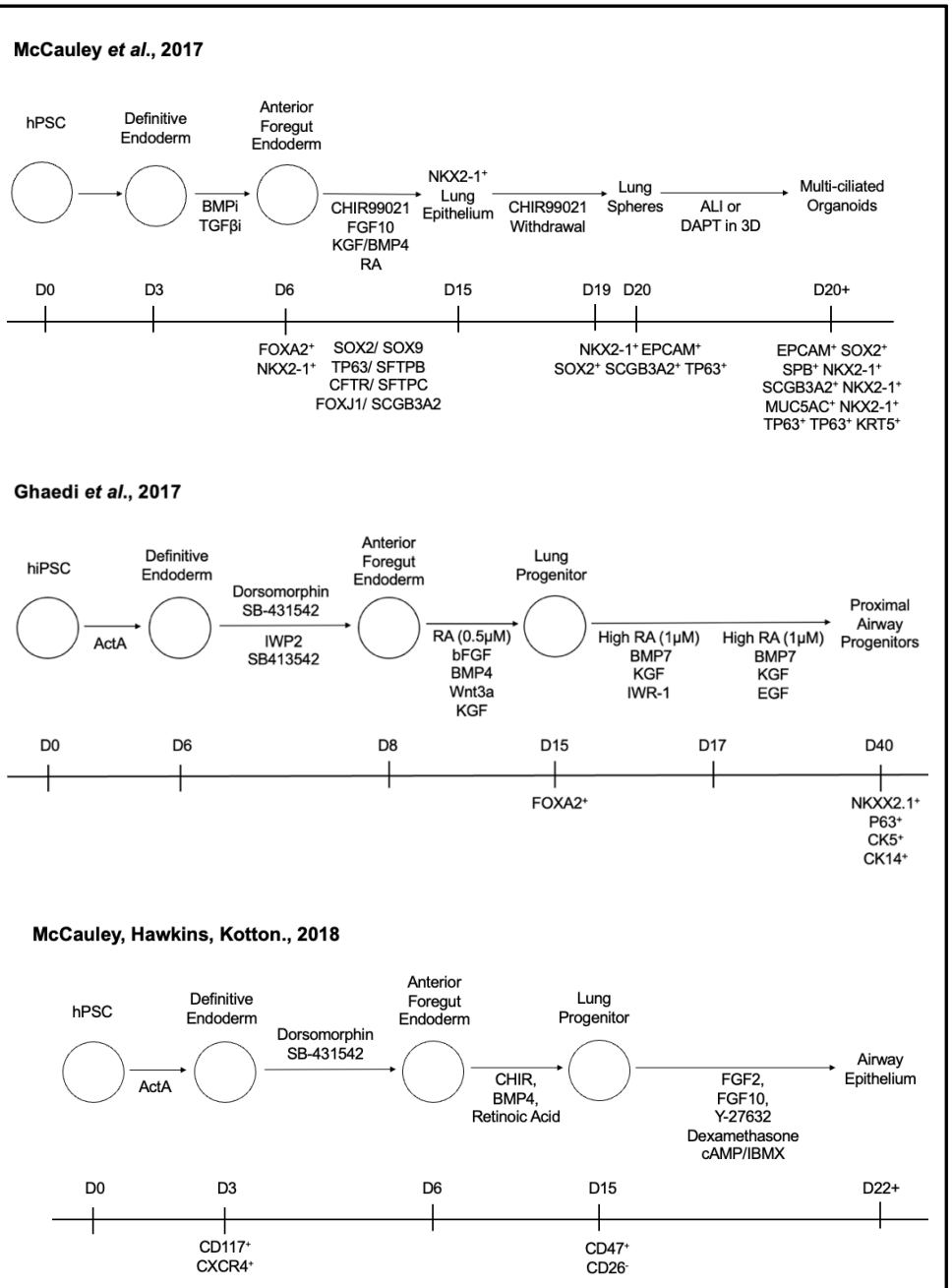


Figure 5-3. Schematic of Published Protocols on Differentiation of iPSC. Protocols by McCauley *et al*, Ghaedi *et al*, McCauley, Hawkins and Kotton are outlined.

5.1.2 Primary Cells-Derived Airway Epithelium

Human bronchial epithelial cells (HBECs) cultured in ALI are currently the gold standard method for modelling the airway epithelium[10]. The differentiated pseudostratified layer possesses greater physiological relevancy to the epithelium *in vivo* compared to undifferentiated cells cultured in submerged conditions[10]. Some studies have shown that HBEC-derived ALI cultures can recapitulate gene expression and secretary profiles of primary cells *in vivo*[613]. Kesimer *et al* compared the protein composition between human tracheobronchial normal induced sputum and HBECs cultured at ALI. A high level (90%) of similarity was observed with the mucins MUC1, MUC4, MUC16, MUC5B and MUC5AC being identified in both samples[613]. Furthermore, Pezzulo *et al*. reported similarity between the gene profile of ALI-HBECS and the airway epithelium *in vivo*[614]. However, a study by Dvorak *et al* found that whilst a strong overall correlation in gene expression existed between HBEC-derived ALI cultures and *in vivo* airway epithelial cells, differences were apparent[615]. Genes associated with immune response and the cytoskeleton had higher expression in vivo cells whereas genes associated with proliferative processes showed elevated expression in ALI cells[615]. This difference could be due to differences in the biology of the cell culture environment[615]. Although ALI models have been shown to recapture the gene expression pattern *in situ* and maintain important metabolic enzymes, the cultures tend to express higher levels of cell cycle and basal-cells related genes[615].

ALI allows investigation of host-virus interactions along with mechanistic studies behind respiratory diseases. Cells from a selected patient population can be studied using primary cells cultured in ALI. For example, characteristics of a distorted airway epithelium are retained in ALI-cultured cells derived from asthmatic patients therefore allowing disease modelling in a physiologically relevant way[616]. Moreover, Ji *et al* to investigate the effects of air pollution nanoparticles using chronic bronchiolitis ALI model[617]. However, disadvantages of the ALI-HEC models include the invasive nature of bronchoscopy which can cause complications such as trauma to the airway, nasopharyngeal, and vocal cord, laryngospasm, bronchospasm, haemorrhage, pneumothorax and infection[618]. Furthermore, variation between

differentiation of the ALI-HBEC model has been observed across different research groups. For example, ALI cultures have been found to resemble the cuboidal morphology of mice epithelium instead of the columnar morphology of the human airway epithelium[591]. Although the ALI-HBEC model provides a useful platform for studying host-virus interactions, alternative methods such as iPSC have been utilised.

5.1.3 Studying Host-Virus Interactions Using Cell Culture Models

Viruses rely on host cells for synthesis of macromolecular machinery, genome replication and assembly of particles[505]. Due to this dependence, viruses have adopted a multitude of mechanisms to exploit host cell functions which often damages the host cells[505]. Investigation of host-virus interaction is important as many vital discoveries surrounding cellular processes were discovered when studying replication cycle of viruses or when using viruses as model systems[505]. Studying the complexities of viral gene expression has resulted in the discovery of host factors that influence host genome replication, protein synthesis and RNA splicing [505]. Host-RSV interaction has mainly been studied using immortalised cell lines[619]. Hallak *et al* investigated RSV infection using laryngeal carcinoma-derived HEp-2 cells[620]. Whereas Krusat and Streckert studied the interaction using cervical carcinoma-derived HeLa cells[621]. However due to the differences in many aspects from *in vivo* airway epithelium, immortalised cell lines may not be the best model to study RSV infection[619]. Few studies have used primary cells to study host-RSV interaction. Study by Johnson *et al* utilised primary, differentiated airway epithelium cultures to elucidate CX3CR1 as a receptor for RSV[505,619].

During viral infection, pattern recognition receptors such as toll-like receptors (TRLs) detect viruses by recognising pathogen-associated molecular patterns[622]. Following the binding of PAMP, TLRs either signals through the TIR-domain containing adapter-inducing IFNβ (TRIF) or the myeloid differentiation factor (MyD88)[622]. Adaptor proteins are recruited, and cellular kinases are activated[622]. This leads to the activation of interferon regulatory factor (IRF) 3 and 7 which initiates production of anti-viral cytokines called interferons (particularly IFNβ, IFNλ) and anti-viral proteins[622]. Anti-viral

proteins, also called restriction factors, such as 2'-5'-oligoadenylate synthetases (OAS) play a vital role in hindering viral replication and limiting viral infection[324]. Furthermore, polymorphisms of OAS1 have been associated with altered cellular function which results in diseases such as Sjögren's syndrome and susceptibility to tuberculosis infection[324]. Therefore, studying the function of OAS1 is important. Interaction between RSV A2 and the host was investigated in this Chapter by assessing the viral load and the expression of RSV receptor CXRCR1, IFNβ, IFNƛ and OAS1 in LOPCK, REBL-PAT, submerged and ALI-HBECSs.

5.2 Aims and Objectives

Overall Aim: To develop iPSC model of the airway epithelium and compare it to gold-standard ALI-HBECS model to investigate host-RSV interactions.

The objectives of the Chapter are:

- To utilise the Wong *et al* protocol to differentiate the iPSC cell line LOPCK and REBL-PAT into lung progenitors and mature airway epithelium.
- To evaluate the differentiation of cells across the stages: 1 (formation of DE), 2 (formation of AFE), 3 (bronchial progenitors), 4 (expansion of bronchial progenitors).
- To assess the utility of the model to study RSV infection by examining how permissive the iPSCs are at different stages of differentiation and examining the changes in the expression of restriction factors OAS1, interferon responses, and CX3CR1 across key stages.

5.3 Methods and Materials

5.3.1 Molecular Biology Methods

Molecular Biology Methods specific to this Chapter such as cDNA synthesis and RT-qPCR are detailed in this section.

5.3.1.1 Markers of iPSC Differentiation

Validation of differentiation at end of each stage is important. Therefore, the cells at each stage of the differentiation are characterised by assaying specific markers. FOXA2, SOX17, NKX2.1, FOXJ1, KRT5 and MUC5AC were assayed for at the end of each stage of differentiation of LOPCK and REBL-PAT iPSCs to see the change in gene expression throughout the stages. The aforementioned markers were also assayed in undifferentiated LOPCK and REBL-PAT iPSCs, submerged and ALI-HBECs. 50 ng of RNA was used for RT-qPCR. The stages of differentiation and function of markers are summarised in table 5-2.

Table 5-2. List of RT-qPCR markers used to assay for different stages of the differentiation. The function of the different markers is outlined in the table below.

Markers	Stages of Differentiation	Function
FOXA2	Stage 1	Formation of DE
SOX17	Stage 1	Formation of DE
NKX2.1	Stage 2	Formation of AFE, lung morphogenesis
FOXJ1	Stage 3, 4, 5	Formation of cilia
KRT5	Stage 3, 4, 5	Formation of basal cells
MUC5AC	Stage 3, 4, 5	Production of MUC5AC protein

5.3.1.2 cDNA Synthesis

Through reverse transcription, complementary DNA (cDNA) was synthesised from an RNA template[623]. Reverse transcriptase (RT), the replicating enzyme of retrovirus, catalyses this reaction[623]. Under well-defined pH and salt conditions with the presence of deoxynucleotides (dNTP), primer complementary to the RNA is extended by RT to produce cDNA[623]. cDNA is used for RT-qPCR.

Protocol

1. Before starting, each component was briefly centrifuged
2. For each reaction, the following components were added to a sterile 0.5-ml Eppendorf tube (Table 5-3)

Table 5-3. Reaction mixture for cDNA synthesis. Reagents and volumes are outlined.

Reagent	Volume (µl)
RNA	1
10 mM dNTP mix	1
Random hexamers	1
DEPC-treated water	7

3. The RNA/primer mixture was incubated for 5 minutes at 65°C and then placed on ice for approximately 1 minute
4. The 2X reaction mix was prepared in a separate tube and with each component being added in the order stated on Table 5-4

Table 5-4. 2X reaction mixture for cDNA synthesis. Reagents and volumes are outlined.

Reagent	Volume (µl)
RNA	2
10 mM dNTP mix	4
Primer	2
DEPC-treated water	1

1. 9 µl of the 2X reaction mix was added to each RNA/primer mixture
2. The reaction mixture was briefly centrifuged and incubated at 25°C for 2 minutes
3. 1 µl of SuperScript™ II RT was added to each tube
4. 1 µl DEPC-treated water instead of the RT was used for the RT-negative control
5. The reaction mixture was incubated at 25°C for 10 minutes and 42°C for 50 minutes
6. The reaction was terminated at 70°C for 15 minutes and chilled on ice
7. The reaction was collected by brief centrifugation
8. 1 µl of RNase H was added to each tube and incubated at 37°C for 20 minutes
9. The reaction mixture was stored at -20°C

5.3.1.3.RT-qPCR

Quantitative polymerase chain reaction (qPCR) enables 'real-time' measurement of gene expression by utilising the use of fluorescently labelled oligonucleotide probes. Detection is only enabled after hybridisation between the probe and its complementary sequence[624]. The quantity of DNA present is proportional to the amount of fluorescence[625]. Following RT-mediated cDNA synthesis, RT-qPCR was used to assay for differentiation markers and for RSV, RSV receptor CX3CR1, interferon response and restriction factor OAS1.

Protocol

1. For each sample, the PCR reaction mix was prepared using the following reagents in a microcentrifuge tube (Table 5-5)

Table 5-5. Components for PCR reaction mix. Reagents and volumes are outlined.

Reagent	Volume (µl)
20x TaqMan® Gene Expression Assay	1
TaqMan® Gene Expression Master Mix	10
cDNA Template	1
RNase-free Water	18

2. The reaction components were mixed by inverting the tube several times
3. The tube was briefly centrifuged
4. 20 µl of PCR reaction mix was transferred to each well of a 96-well reaction plate
5. The plate was sealed and briefly centrifuge the plate
6. The plate was loaded onto a thermocycler
7. The thermocycling conditions are stated on Table 5-6

Table 5-6. Thermocycling conditions used for RT-qPCR. Temperature and duration for each stage are outlined.

Stage	Temperature (°C)	Time
Hold	50	2 minutes
Hold	95	10 minutes
40 Cycle	95	15 seconds
40 Cycle	60	1 minute

5.3.1.4.RT-qPCR Data Analysis

For data analysis, delta cycle threshold (ΔCt) was used. Ct value is the number of cycles required for the target gene to exceed a threshold value. The Ct values for the technical replicate for each sample was averaged. Next, the delta Ct (ΔCt) which is the difference between Ct values of the gene of interest and housekeeping gene was calculated. Housekeeping genes are important for maintaining basic cellular function and are expressed in all cell types[626]. They are used as a method for normalisation[626]. 18S ribosomal RNA (18S rRNA) was used as the housekeeping gene because the gene was expressed in the key stages of iPSC differentiation, submerged and ALI-HBECs. ΔCt was calculated using the following formula:

$$\Delta Ct = Ct\ (\text{Gene of Interest}) - Ct\ (\text{Housekeeping Gene})$$

5.3.2 Cell Culture Methods

This section outlines the cell culture methods used for this Chapter.

5.3.2.1 iPSC

This section details the iPSC cell line, reagents and protocols used during the thesis.

5.3.2.1.1 REBL-PAT Cell Line

REBL-PAT cell line was one of the iPSC cell line used for this Chapter. The cell line is derived from a skin punch biopsy of dermal fibroblast taken from a male subject[627]. CytoTune™-IPS 2.0 Sendai Reprogramming Kit was used to transduce the cell line[627]. The cell line was kindly provided by the Hannan Lab Group.

5.3.2.1.2 LOPCK Cell Line

LOPCK cells was another iPSC cell line used for the Chapter. The cell line was kindly provided by the Hannan Lab Group.

5.3.2.1.3 Coating of Cell Culture Dishes with Matrigel

Prior to culturing cells, cell culture dishes were coated with Matrigel overnight. Matrigel is an extract of basement membrane proteins derived from Engelbreth-Holm-Swarm murine tumour[628]. The extract is reconstituted into a solid gel at 37°C[629]. Matrigel supports differentiation and cell morphogenesis[628].

Protocol

1. For T25 flask, 73 µl of Matrigel (11.9 mg/ml) was suspended in 10 ml of DMEM F-12 media
2. 5 ml of the Matrigel-DMEM F-12 solution was placed in a T25 flask
3. The T25 flask was incubated overnight at 37°C and 5% CO_2 in an incubator
4. For 6-well plate, 1.5 ml of Matrigel solution was used to coat each well

5.3.2.1.4 Maintenance of iPSC

iPSCs were maintained daily using Homebrew Essential 8 media (HB E8). Following reagents were used to make up the media (Table 5-7).

Table 5-7. Volume of reagents was used to make the HB E8 media for maintainence.

Reagent	Volume (µl)
Insulin	100
Small Molecule	75
Transferrin	9
Penicillin-streptomycin	500

Protocol

1. 5 µl of the HB E8 media in a 5 ml universal tube was warmed in water bath
2. 5 µl of growth factors FGF2 (1:1000) and TGF-β (1:1000) was added to HB E8 media
3. The flask was incubated at 37°C and 5% CO_2 in a cell culture incubator
4. Media was changed daily

5.3.2.1.5 Adaption of iPSC to Single-cell Passaging

Cells can have reduced mitotic index if confluent cells are not passaged[630]. This can lead to cell death. Therefore, the cells are passaged routinely. First the cells are detached from the cell culture vessel through trypsinisation. TrypLE, a recombinant enzyme with trypsin-like activity, was used to trypsinise the cells[631]. The cell suspension is then replated into fresh culture. The method for passaging iPSC is detailed below.

Protocol

1. 5 ml of the HB E8 media was aliquoted into a universal tube
2. The media was warmed to 37°C in a water bath
3. The following growth factors was added to the media (Table 5)
4. Matrigel in the Matrigel-coated T25 flask was washed twice with PBS and aspirated
5. Spent media from the T25 flask containing iPSC was aspirated
6. The iPSC containing-T25 flask was washed twice with 5 ml of PBS and aspirated
7. The iPSCs were incubated with 3 ml of TrpLE for 4 minutes at 37°C and 5% CO_2
8. After 4 minutes, TrpLE was immediately aspirated
9. The iPSCs were flushed against the wall 3 times using 5 ml of HB E8 media without the growth factors
10. The cells were transferred and passaged in a 1:20 ratio with 5 ml of HB E8 media with growth factors onto the Matrigel-coated T25 flask
11. The iPSCs were incubated in a cell culture hood at 37°C and 5% CO_2

5.3.2.1.6 Differentiation of iPSC into Lung Cells

Using Wong *et al* protocol, iPSC cell line LOPCK and REBL-PAT were differentiated through the following stages: Stage 1 definitive endoderm, Stage 2 anterior foregut epithelium, Stage 3 bronchial progenitors, Stage 4 expansion of bronchial progenitors and Stage 5 maturation and polarisation of lung epithelia[610]. Following establishment of DE, the cells are directed to AFE using high concentration of FGF2 and SHH[610]. Cells are driven to a lung fate at high concentrations of FGF2, whereas intermediate concentration drives pancreatic lineage, and low concentration promotes hepatic development[632]. The distal lung bud shows high expression of SHH where it mediates epithelial proliferation and mesenchymal cell growth during morphogenesis of lung branches[633]. Furthermore, the drive towards the pancreatic fate is suppressed by SHH[634].

The cells are then exposed to FGF7, FGF10, FGF18 and BMP4 to drive differentiation into immature lung cells[610]. FGF7 promotes epithelial proliferation and distinct growth patterns of the airway epithelium[605]. Located in the lung mesenchyme, FGF10 is vital for epithelial proliferation and commitment to the lung lineage during embryonic and post-natal development[635]. FGF10 is also involved in the outgrowth of lung bud, organogenesis, and protection and regeneration of the lung epithelium following injuries[635]. FGF18 also plays an essential role in lung organogenesis[636]. The expression of FGF18 is elevated during the pseudo-glandular stage of lung development when maximal branching occurs[636]. Whereas BMP4 mediates epithelial cell differentiation and morphogenesis in developing lungs[637]. To induce maturation and differentiation into a pseudostratified airway epithelium complete with basal, ciliated and mucus-producing goblet cells, the cells are cultured in ALI for five weeks[610]. Summary of the growth factors used for the differentiation is outlined on Table 5-8.

Table 5-8. List of growth factors used during different stages of the differentiation. The function of the different growth factors is outlined in the table below.

Growth Factor	Stages of Differentiation	Function
FGF2	Stage 2	Drives lung fate
SHH	Stage 2	Morphogenesis of lung branches
FGF7	Stage 3	Epithelial proliferation, induces expression of SP-A and SP-B
FGF10	Stage 3	Epithelial proliferation, drives lung fate
BMP4	Stage 3	Epithelial differentiation and morphogenesis of lung branches
FGF18	Stage 3 and 4	Lung organogenesis

Protocol

Stage 1: Differentiation to Definitive Endoderm

Prior to differentiation, 6-well plate were coated using the protocol outlined on section 2.2.2.3.

Day -1

1. 55,000 cells were seeded per well in a 6-well plate as outlined on section 2.2.2.3.
2. 1.5 ml of the cell suspension was added per well of the 6-well plate
3. The cells were incubated at 37°C and 5% CO_2 in a cell culture incubator

Day 0

1. 50 ml of RPMI basal media was prepared using the reagents from Table 5-9

Table 5-9. Volume and concentration of reagents used to make the RPMI basal media.

Reagents	Volume (ml)	Concentration
B27	1.00	0.10%
Penicillin-streptomycin	0.50	0.10%
Non-essential amino acids	0.50	0.10%

2. Spent media was aspirated from the 6-well plate
3. 9 µl of ActA and Wnt was added to 9 ml of RPMI basal media
4. 1.5 ml of RPMI basal media was added per well of the 6-well plate
5. The cells were incubated at 37°C and 5% CO_2 in a cell culture incubator

Day 1

1. Media was replaced with 1.5 ml of RPMI basal media supplemented with ActA and Wnt per well

Day 2

1. Media was replaced with 1.5 ml of RPMI basal media supplemented with ActA and Wnt per well
2. For the coating of Transwell insert, 1 ml of human placenta collagen was added to 9 ml of deionised water to make a working concentration of 60 µg/ml
3. 150 µl of collagen was added per Transwell insert and left open overnight in a sterile cell culture cabinet

Day 3

1. Cells were lysed as outlined on section 2.1.2 to assay for DE markers by RT-qPCR

Stage 2: Differentiation to the Anterior Foregut Endoderm

Day 3

1. The collagen-coated Transwell inserts were washed three times with DMEM/F12
2. The transwell inserts were air dried for at least 1 hour in the sterile cell culture incubator
3. The differentiation basal medium was prepared by adding the following reagents to 50 ml of KnockOut™ DMEM media (Table 5-10).

Table 5-10. volume and concentration of regents used to make the differentiation basal medium.

Reagent	Volume (ml)	Concentration
KnockOut™ Serum Replacement	5.00	10%
Penicillin-streptomycin	0.30	1%
GlutaMAX™	0.60	2 mM
1-Thioglycerol (MTG)	0.15	0.15 mM
Non-essential Amino Acid	0.30	1 mM

4. Spent medium was aspirated from the 6-well plate
5. Cells were washed with PBS twice
6. Cells were incubated with 300 µl of TrypLE per well for 5 minutes at 37°C and 5% CO_2
7. 1 ml of basal differentiation medium was added per well
8. Cells were transferred to a 15 ml conical tube and centrifuged at 400g at 25°C temperature for 5 minutes
9. The supernatant was aspirated, and cell were suspended in 1 ml of differentiation basal medium
10. Anterior foregut endoderm (AFE) medium was prepared by adding 500 ng/ml of FGF2, 50 ng/ml SHH and 50 ng/ml Rock inhibitor to prewarmed differentiation basal medium
11. ~5 × 10^5 cells with AFE medium were seeded to the apical side of the Transwell inserts
12. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator

Day 4

1. Spent media was aspirated and 0.25 ml of AFE media supplemented with 500 ng/ml of FGF2 and 50 ng/ml SHH was added to the apical side of the Transwell

Day 5

1. Spent media was aspirated and 0.25 ml of AFE media supplemented with 500 ng/ml of FGF2 and 50 ng/ml SHH was added to the apical side of the Transwell
2. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 7

1. Spent media was aspirated and 0.25 ml of AFE media supplemented with 500 ng/ml of FGF2 and 50 ng/ml SHH was added to the apical side of the Transwell
2. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 24 hours

Stage 3a: Differentiation to Lung Progenitors

Day 8

1. Stage 3a medium was prepared by combining add 50 ng/ml of FGF7, 50 ng/ml of FGF10 and 5 ng/ml BMP4 to differentiation basal medium
2. Spent media was aspirated and 0.25 ml of Stage 3a media was added to the apical side of the Transwell
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 24 hours

Day 10

1. Spent media was aspirated
2. 0.25 ml of AFE media supplemented with 50 ng/ml of FGF7, 50 ng/ml of FGF10 and 5 ng/ml BMP4 was added to the apical side of the Transwell
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 12

1. Spent media was aspirated
2. 0.25 ml of AFE media supplemented with 50 ng/ml of FGF7, 50 ng/ml of FGF10 and 5 ng/ml BMP4 was added to the apical side of the Transwell
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Stage 3b: Differentiation to Immature Lung Cells

Day 13

1. Stage 3b medium was made by adding 10 ng/ml of FGF7, 10 ng/ml of FGF10 and 10 ng/ml of FGF18 to the differentiation basal medium
2. 0.25 ml of stage 3b medium was added to the apical side of the Transwell inserts

3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 15

1. Stage 3b medium was aspirated from the cells
2. 0.25 ml of stage 3b medium supplemented with 10 ng/ml of FGF7, 10 ng/ml of FGF10 and 10 ng/ml of FGF18 was added to the apical side of the Transwell inserts
3. Cells were incubated at 37 °C and 5% CO_2 in a cell culture incubator for 48 hours

Day 17

1. Stage 3b medium was aspirated from the cells
2. 0.25 ml of stage 3b medium supplemented with 10 ng/ml of FGF7, 10 ng/ml of FGF10 and 10 ng/ml of FGF18 was added to the apical side of the Transwell inserts
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Stage 4: Expansion of Immature Lung Cells

Day 18

1. Lonza BEGM™ Bronchial Epithelial Cell Growth Medium (BEGM) BulletKit™ was prepared by combining the supplements to 500 ml of BEBM® Basal Medium (Table 5-11)

Table 5-11. Volume and concentration of reagents used to make the BEGM™ media.

Reagent	Volume (ml)	Concentration
Bovine Pituitary Extract (BPE)	2.00	52.00 µg/ml
Insulin	0.50	5.00 µg/ml
Hydrocortisone	0.50	0.50 µg/ml
Gentamicin Sulfate-Amphotericin (GA-1000)	0.50	Gentamicin: 30 µg/ml Amphotericin: 15 µg/ml
Retinoic Acid	0.50	0.10 ng/ml
Transferrin	0.50	10.00 µg/ml
Triiodothyronine	0.50	6.50 ng/ml
Epinephrine	0.50	0.50 µg/ml
Human Epidermal Growth Factor (hEGF)	0.50	0.50 ng/ml

2. 10 ng/ml of FGF18 was added to the BEGM™ medium
4. Spent 3b media was aspirated and 0.25 ml of BEGM™ media with 10 ng/ml of FGF18 was added to the apical side of the Transwell
5. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 20

1. Spent media was aspirated and 0.25 ml of BEGM™ media with 10 ng/ml of FGF18 was added to the apical side of the Transwell
2. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 22

1. Spent media was aspirated and 0.25 ml of BEGM™ media with 10 ng/ml of FGF18 was added to the apical side of the Transwell
2. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 24 hours

Day 23

1. Using the B-ALI™ Bronchial Air-Liquid Interface Medium BulletKit™, the following reagents was combined to prepare 250 ml of B-ALI™ Growth Medium (Table 5-12)

Table 5-12. Volume and concentration of reagents used to make B-ALI™ Growth Medium.

Reagent	Volume (ml)	Concentration
Bovine Pituitary Extract (BPE)	1.00	52.00 µg/ml
Insulin	0.25	5.00 µg/ml
Hydrocortisone	0.25	0.50 µg/ml
Gentamicin Sulfate-Amphotericin (GA-1000)	0.25	Gentamicin: 30 µg/ml Amphotericin: 15 µg/ml
Retinoic Acid	0.25	0.10 ng/ml
Transferrin	0.25	10.00 µg/ml
Triiodothyronine	0.25	6.50 ng/ml
Epinephrine	0.25	0.50 µg/ml
Human Epidermal Growth Factor (hEGF)	0.25	0.50 ng/ml

1. Spent BEGM™ growth medium was aspirated
2. 0.25 ml of B-ALI™ growth medium was added to the apical side of the Transwell inserts

3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 25

1. Spent BEGM™ growth medium was aspirated
2. 0.25 ml of B-ALI™ growth medium was added to the apical side of the Transwell inserts
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours

Day 27

1. Spent BEGM™ growth medium was aspirated
2. 0.25 ml of B-ALI™ growth medium was added to the apical side of the Transwell inserts
3. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 24 hours

Stage 5: Maturation and Polarization of the Epithelium

Day 28

1. Spent B-ALI™ growth medium was aspirated
2. 500 ml of B-ALI™ Differentiation Medium was prepared by combining the following reagents (Table 5-13)

Table 5-13. Volume and concentration of reagents used to make B-ALI™ Differentiation Medium

Reagent	Volume (ml)	Concentration
Bovine Pituitary Extract (BPE)	2.00	52.00 µg/ml
B-ALI-Inducer	0.50	0.03 µg/ml
Insulin	0.50	5.00 µg/ml
Hydrocortisone	0.50	0.50 µg/ml
Gentamicin Sulfate-Amphotericin (GA-1000)	0.50	Gentamicin: 30 µg/ml Amphotericin: 15 µg/ml
Retinoic Acid	0.50	0.10 ng/ml
Transferrin	0.50	10.00 µg/ml
Triiodothyronine	0.50	6.50 ng/ml
Epinephrine	0.50	0.50 µg/ml
Human Epidermal Growth Factor (hEGF)	0.50	0.50 ng/ml

1. 0.50 ml of B-ALI differentiation medium was added to the basolateral side of the Transwell inserts
2. Cells were incubated at 37°C and 5% CO_2 in a cell culture incubator for 48 hours
3. Steps 1-3 of Stage 5 was repeated up to 5 weeks with media change every 48 hours

5.3.2.2 HBECs

5.3.2.2.1 Maintenance of HBECs using PneumaCult™-Ex Plus Medium

Prior to ALI, HBEC D195307 was cultured in submerged conditions using PneumaCult™-Ex Plus Medium as detailed on Chapter 2 Section 2.2.3.3. The PneumaCult™-Ex Plus Medium is a bovine pituitary extract (BPE) and serum-free medium which enables more expansion of HBECs at each passage. The medium possesses a better differentiation potential making it an ideal medium for forming pseudostratified epithelium at ALI.

Protocol

1. PneumaCult™-Ex Plus 50X Supplement was thawed at 25°C and gently inverted
2. 10 ml of PneumaCult™-Ex Plus 50X Supplement and 0.5 ml of Hydrocortisone Stock Solution was added to 490 ml of PneumaCult™-Ex Plus Basal Medium and mixed
3. 11 ml of reconstituted PneumaCult™-Ex Plus Basal Medium was added to T75 flaks
4. The cells were cultured at 37°C and 5% CO_2 in a cell culture incubator

5.3.2.2.2 Air Liquid Interface

HBECs are cultured in ALI conditions to form a pseudostratified mucociliary layer with specialised cells such as basal, ciliated and goblet cells. Upon confluency, HBEC D195307 was passaged and seeded as detailed on Chapter 2 Section 2.2.3.3. HBEC D195307 were cultured in PneumaCultTM-ALI^{TM} in ALI conditions.

Protocol

1. PneumaCultTM-ALI^{TM} 10X supplement was defrosted overnight at 4°C and gently inverted
2. 50 ml of the 10X supplement was added to 450 ml of PneumaCultTM-ALI^{TM} Basal Medium and mixed thoroughly to make up 500 ml of PneumaCultTM-ALI^{TM} Complete Base Medium
3. 20 ml of PneumaCultTM-ALI^{TM} Complete Base Medium as supplemented using the reagents from Table 5-14

Table 5-14. Reagents for PneumaCultTM-ALI^{TM} Complete Base Medium.

Reagent	Volume
PneumaCultTM-ALI Maintenance Supplement	200 µl
Heparin	40 µl
Hydrocortisone	100 µl
Penicillin-streptomycin	200 µl

4. 500 µl of PneumaCultTM-ALI^{TM} Complete Base Medium was placed in the basolateral chamber of the Transwell insert
5. The ALI inserts were incubated at 37ºC and 5% CO_2 in a cell culture incubator

5.3.2.2.3 Trans-epithelial Electrical Resistance Readings

Trans-epithelial electrical resistance (TEER) is a routinely used quantitative method to measure integrity of tight junction[624]. Electrical resistance is measured across a cellular monolayer to enable a reliable and sensitive method to assess barrier integrity and permeability[624]. The EVOM Epithelial Voltohmmeter (World Precision Instruments) was used to conduct TEERs on day 7, 14 and 21 of ALI.

Protocol

1. 200 µl of PneumaCultTM-ALI^{TM} Complete Base Medium was placed into the apical surface and 500 µl into the basolateral chamber of the Transwell inserts including the blank
2. The inserts were incubated for 20 minutes at 37ºC and 5% CO_2
3. The electrode was calibrated by exposing it to 70% IMS, PBS and then ALI media
4. The electrode was then placed on the insert
5. Three readings were taken for each insert and the average was calculated

5.3.3 Viral Methods

This section details the viral methods used in this Chapter.

5.3.3.1 RSV-A2 GFP

RSV-A2 GFP (rgRSV) was kindly provided by Professor Colin Bingle[369]. A full-length RSV cDNA clone, MP224, was constructed to contain GFP by inserting a BstXI fragment consisting of the gene start, NS1 untranslated region, the Green Lantern Protein (Life technologies, Inc.) and the L gene end in this particular order[369]. To generate the BstXI fragment, the BstXI/BamHI fragment from the C41-GFP minigenome was transferred into another GFP minigenome called the MP129 to generate MP166[369]. The XhoI/NcoI fragment from MP166 was transferred into MP90, a bipartite minigenome which contains the GFP gene and is flanked by BstXI sites and luciferase gene, to generate MP169[369]. The BstXI fragment was transferred from MP169 into a full-length RSV cDNA clone, D46, to create MP224[369]. MP224 was rescued by transfecting it with four plasmids which expressed the N, P, L and M2-1 support proteins into HEp-2 cells[369]. T7 RNA polymerase from the recombinant vaccinia virus MVA-T7 was used to drive transcription from the plasmids. Infectious viruses that were released were amplified in HEp-2 cells[369].

5.3.3.2 Infection of iPSCs and HBECs with RSV-A2

iPSC cell line LOPCK and REBL-PAT and D195307 HBEC were infected using the protocol outlined on Chapter 2 Methods and Materials Section 2.3.3. Cells were infected with RSV-A2 for 1 hour and replaced with fresh media. The cell culture media was removed after 48 hours. The cells were then lysed with Buffer RLT and stored at -80℃ for RT-qPCR.

5.3.4 Immunohistochemistry Methods

The immunohistochemistry (IHC) methods are detailed in this section.

5.3.4.1 Embedding ALI-HBECS culture with 1% Low Melting Agarose Gel

Tissue processing is a technique where a sample is embedded in formalin-fixed paraffin to facilitate sectioning before IHC[638]. The process consists of dehydration, clearing and infiltration[638]. The sample is dehydrated to make aqueous formalin-fixed tissues miscible with non-aqueous paraffin wax[638]. Since dehydration agents are not miscible with paraffin, clearing agents (which are miscible with paraffin and dehydrating agent) are added[638]. Infiltration using paraffin wax is conducted to maintain the architecture of the sample during sectioning[638]. Before tissue processing, the ALI-HBEC cultures were fixed with PFA and embedded with 1% agarose gel.

Protocol

1. 1 gm of agarose was added to 100ml of distilled water in a 250ml glass bottle
2. The mixture was heated in the microwave for 2 minutes and mixed after one minute
3. The ALI-HBEC cultures were fixed with 4% PFA for 2 hours
4. The PFA was discarded, and the ALI-HBEC cultures were washed three times with PBS
5. Plastic cassettes were chilled on ice bath
6. Drops of agarose gel was added to the cassette
7. The membrane from the ALI-HBEC Transwell inserts were cut using disposable blades and transferred onto the gel with forceps
8. Drops of gel were immediately dropped onto the samples and left on ice for 5 minutes
9. The cassettes were kept in the 4°C fridge room for 24 hours
10. After 24 hours, the ALI-HBEC cultures were taken for tissue processing

5.3.4.2 Use of Microtome for Sectioning

Following tissue processing, the paraffin embedded ALI-HBEC cultures were sectioned using the microtome.

Protocol

1. Water bath was filled with distilled water and heated to 40-45°C
2. Disposable blade was clamped onto the microtome by folding the blade guard and the right clamping lever downwards
3. The blade was slid in from the side in-between the knife holder and the clamp mount
4. The blade was clamped by folding the right clamping lever and the blade guard upwards
5. The embedded block was inserted into the specimen clamp
6. To section the specimen, the blade guard was folded downwards and handwheel locking mechanism was released until the block was sufficiently trimmed
7. Sections were picked up using a tweezer a floated onto the surface of the water bath to flatten out
8. Sections were picked up from the water bath using the microscope slide
9. Slides were labelled with pencil and dried overnight and stored in room temperature for use

5.3.4.3 Staining

The slides were stained using haematoxylin and eosin staining (H&E) and Alcian blue periodic acid-Schiff (PAS) staining.

Protocol

1) The microscope slide with the ALI sections were placed onto the slide holders
2) The individual reagents were added to individual glass boxes and containers
3) The slide holder was placed into the reagents in the glass boxes and containers for a time period as described below:

Table 5-15. Slide holder containing the microscope slide with ALI sections were placed in the reagents for a time period as stated below.

Reagents	Time (Minutes)
Xylene	5
100% Ethanol	5
100 % Ethanol	5
95% Ethanol	5
95% Ethanol	5
Wash with H_2O	2
3% Acetic Acid	3
Alcian Blue	30
Wash with H_2O	5
0.5% Periodic Acid	10
Wash with H_2O	5
Schiff Reagent	20
Wash with H_2O	10
95% ethanol	5
95% ethanol	5
100% ethanol	5
100% ethanol	5
Xylene	5

1) The insert was mounted by placing two drops of DPX Mountant onto the ALI section
2) Coverslip was placed on top of the section and allowed to dry

5.3.5 Proteomic Methods

The proteomic methods are detailed in this section.

5.3.5.1 Flow Cytometry

Flow cytometry is a proteomic technique that delivers rapid multi-parametric analysis of single cells[639]. Lasers act as light sources to produce fluorescent and scattered light signals which are read by detectors and converted into electronic signals[639]. Populations of cells can be analysed based on their light scattering and fluorescent properties[639]. Flow cytometry was used to measure the percentage of RSV-infected cells by detecting the green fluorescent protein (GFP). Using the protocol outlined on Chapter 2 Methods and Materials Section 2.3.3, iPSC cell line LOPCK were infected with RSV-A2 for 1 hour. Media was replaced with fresh media after 1 hour and the cell culture media was removed after 48 hours. The cells were fixed in PFA for 15 minutes and stored at -80℃ for flow cytometry. The ID7000™ Spectral Cell Analyzer was used for flow cytometry. Flow Cytometry Standard (FCS) files were exported from the ID7000, and the Kaluza software was used for analysis.

5.3.6 Statistical Analysis

5.3.6.1 Analysis of Variance

To identify variance in the means between two or more groups, statistical tests were conducted. One-way ANOVA was conducted for analysis of variance between 3 or more groups. Repeated measures one-way ANOVA was used to compare variance in dependent variables. A two-way ANOVA was used when comparing variance between groups where there were two ways of classifying the data. ANOVA is applied to normally distributed data. Shapiro Wilk test was used to test whether the data was normally distributed. Tukey's Post-hoc (multiple comparisons) test was used to examine the effect of each group on variation. GraphPad Prism 10 was used to carry out the statistical analysis.

5.4 Results

5.4.1 LOPCK and REBL-PAT iPSCs Were Differentiated Into Stage 4 Bronchial Lung Progenitor Cells

The differentiation protocol by Wong *et al* was used to differentiate the iPSC cell line LOPCK and REBL-PAT (n=3). Through the addition of growth factors at an optimum concentration in a carefully timed manner, the iPSCs were differentiated into stage 4 bronchial progenitor cells. iPSCs were passaged prior to differentiation using TrypLE and ROCK inhibitor. Initiation of stage 1 involved culturing the iPSCs in RPMI media supplemented with ActA and Wnt. iPSCs were driven to stage 2 differentiation using Knock-Out DMEM media with FGF2 and SHH. To drive the cells to a lung cell fate, iPSCs were exposed to FGF7, FGF10, BMP4 and FGF18. The bronchial lung progenitor cells were expanded during stage 4 using Lonza BEBM media supplemented with FGF18. RT-qPCR was used to assay for markers at the end of each stage. Moreover, the markers were also assayed in submerged and ALI-differentiated HBECs. At the start of stage 1, the cells form colonies. Stage 1 LOPCK cells exhibited a petal-shape morphology with tight cell-to-cell contact being formed (Figure 5-4). Tight cell-to-cell contact is also observed in stage 1 REBL-PAT cells. As the cells proliferate, they form a monolayer at the end of stage 1. The morphology of LOPCK and REBL-PAT cells appeared as a monolayer during stage 2. Stage 3 cells displayed an epithelial monolayer appearance. At stage 4, cuboidal morphology among cells was observed.

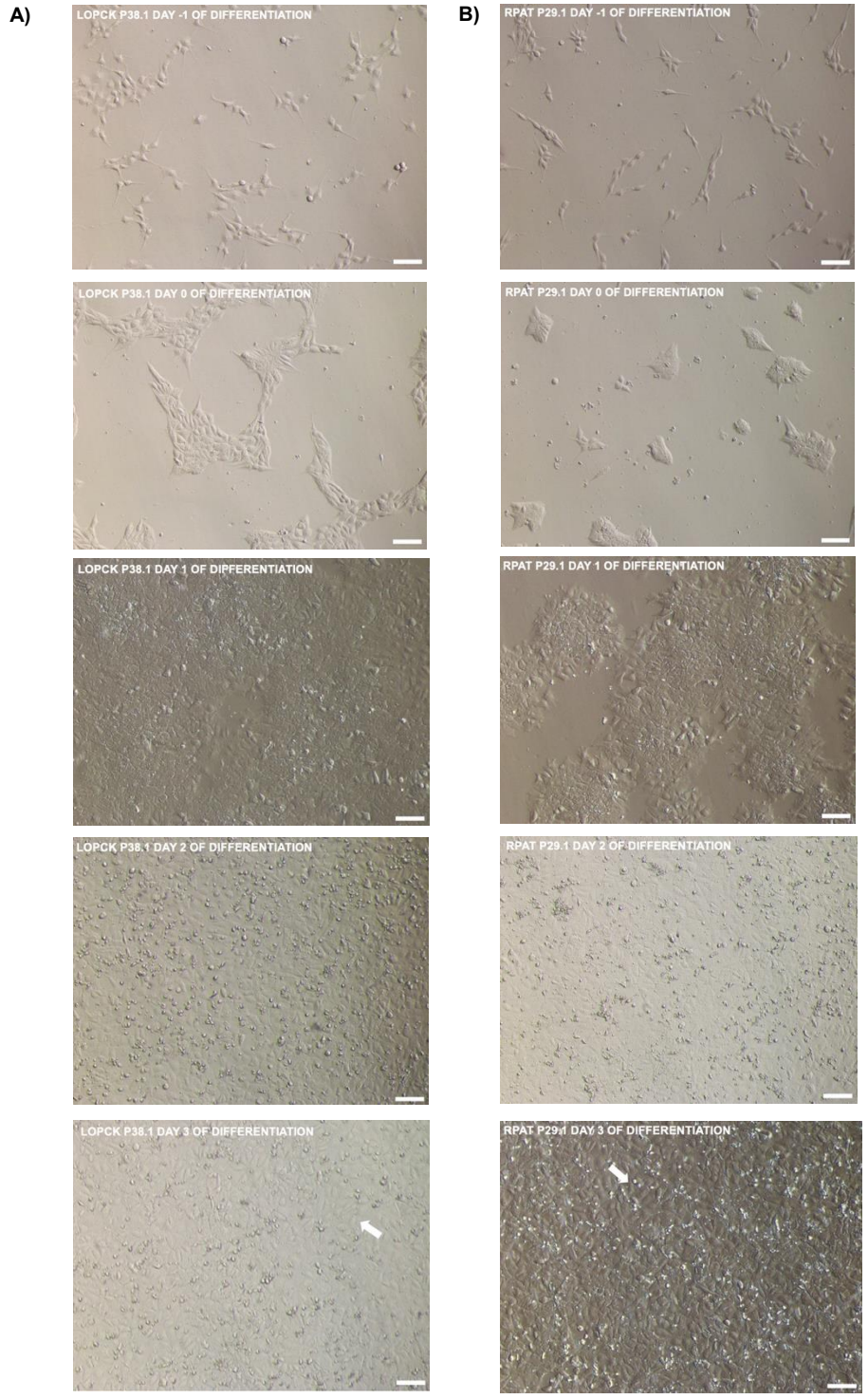


Figure 5-4. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 1 Definitive Endoderm. A) LOPCK P38.1. B) REBL-PAT P29.1. iPSCs were cultured in RPMI media supplemented with Act and Wnt. Scale bar: 100 µm. n = 3. Petal shape morphology (white arrows) can be observed for both cell type at the end of stage 1.

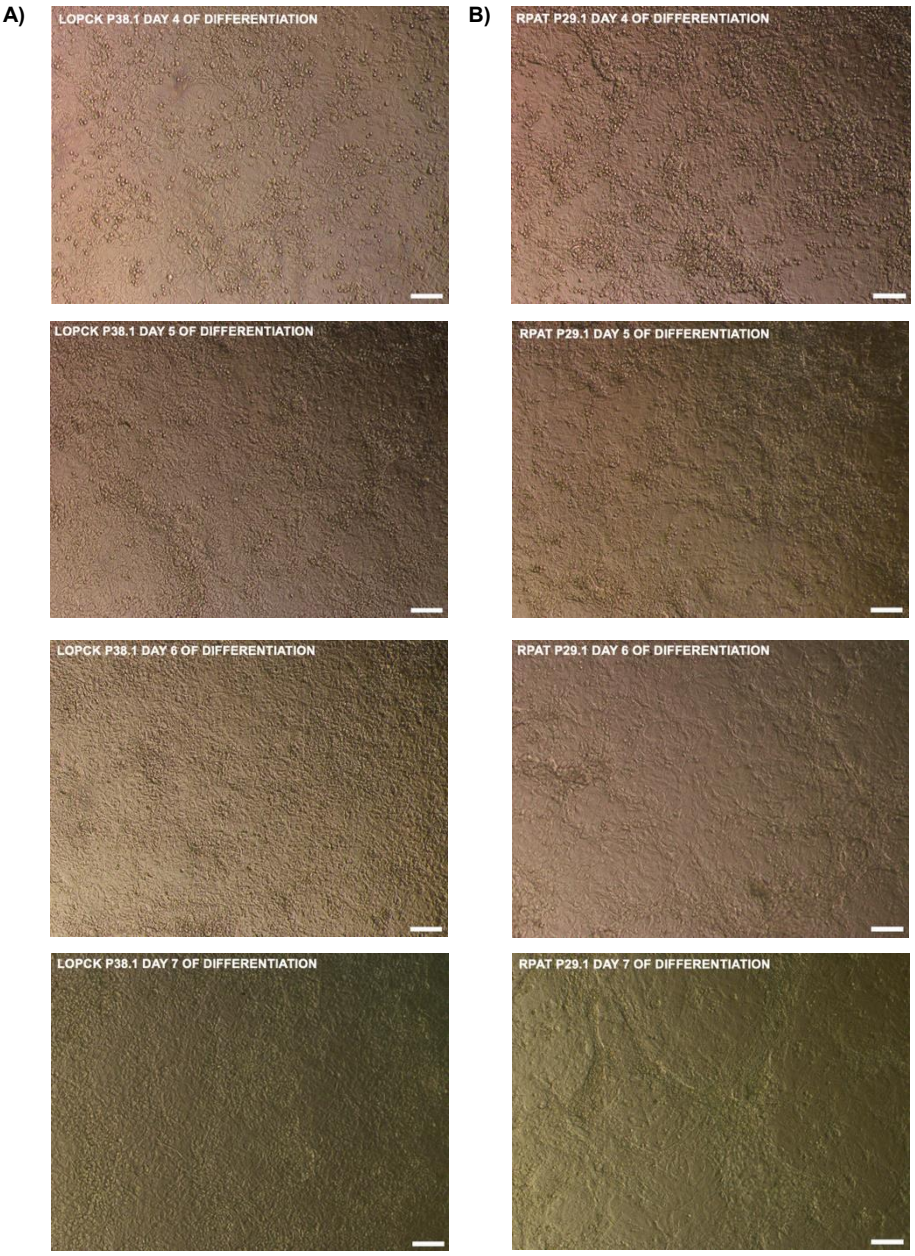


Figure 5-5. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 2 Anterior Foregut Epithelium. A) LOPCK P38.1. B) REBL-PAT P29.1. Cells were cultured in RPMI media with FGF2 and SHH to induce the formation of stage 2 AFE. Scale bar: 100 µm. n = 3. Cells formed a monolayer during stage 2.

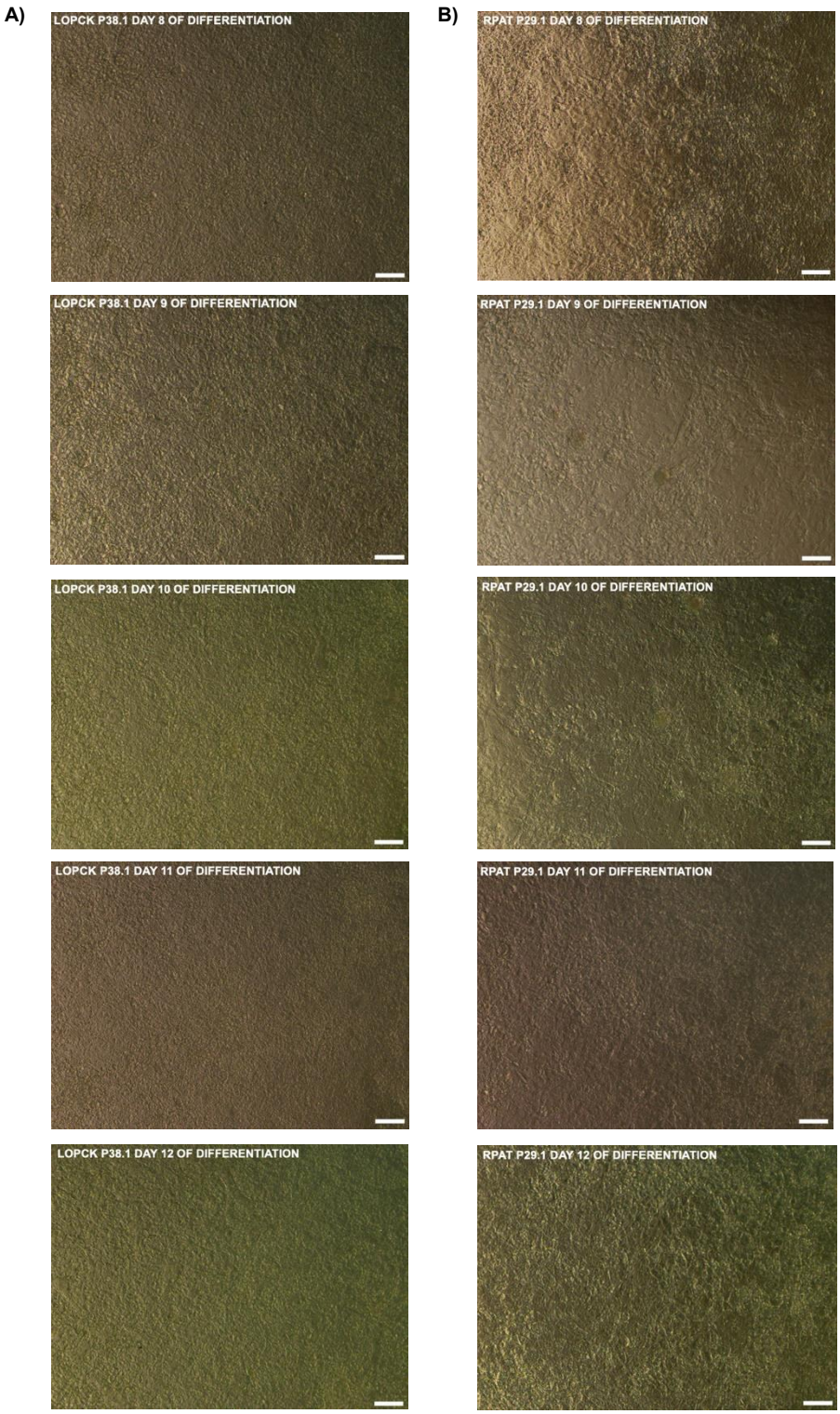


Figure 5-6. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 3a Bronchial Lung Progenitors. A) LOPCK P38.1. B) REBL-PAT P29.1. Cells were cultured in RPMI media with FGF7, FGF10 and BMP4. Scale bar: 100 µm. n = 3.

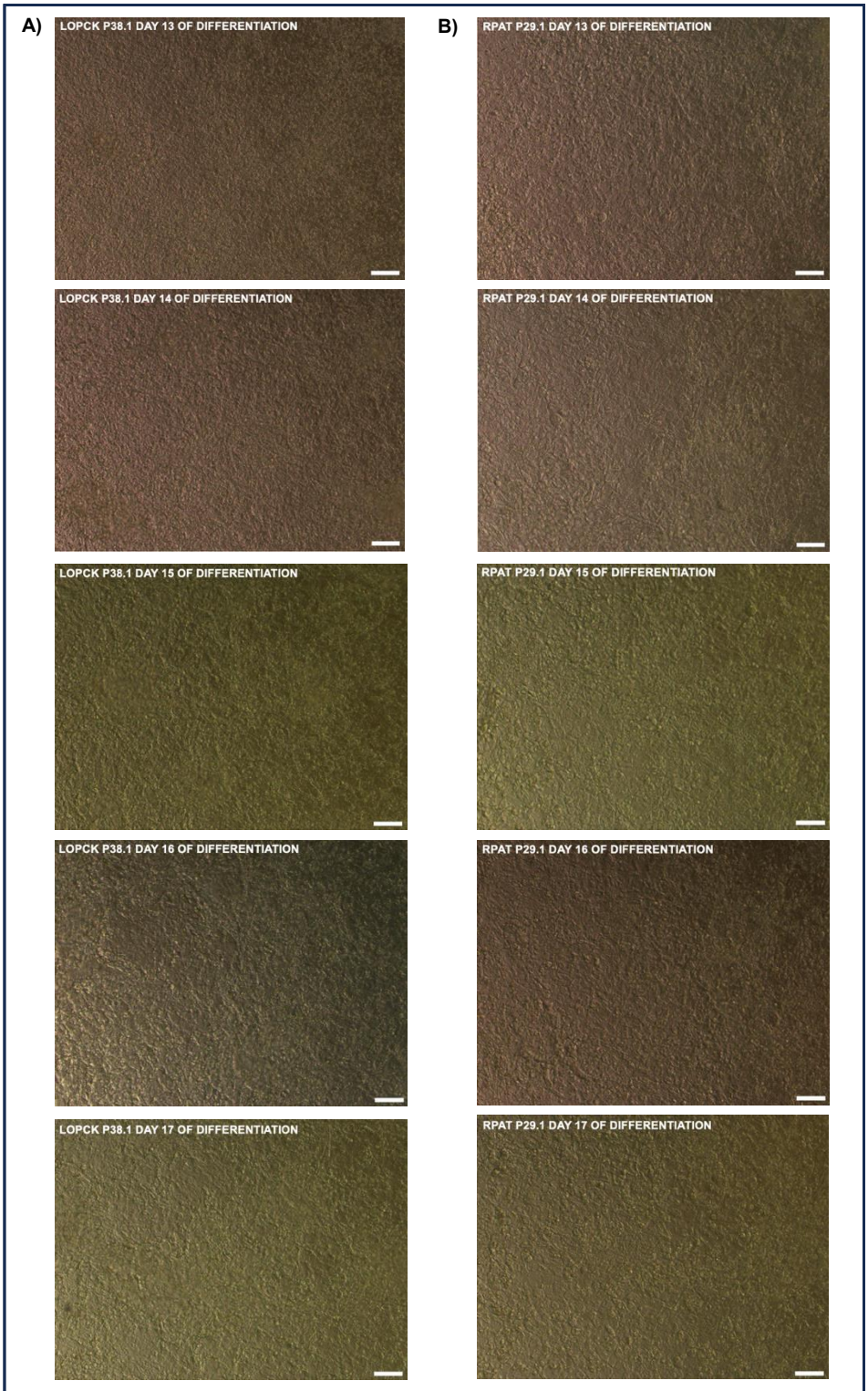


Figure 5-7. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 3b Bronchial Lung Progenitors. A) LOPCK P38.1. B) REBL-PAT P29.1. Cells were cultured in RPMI media with FGF7, FGF10 and FGF18. Scale bar: 100 µm. n = 3.

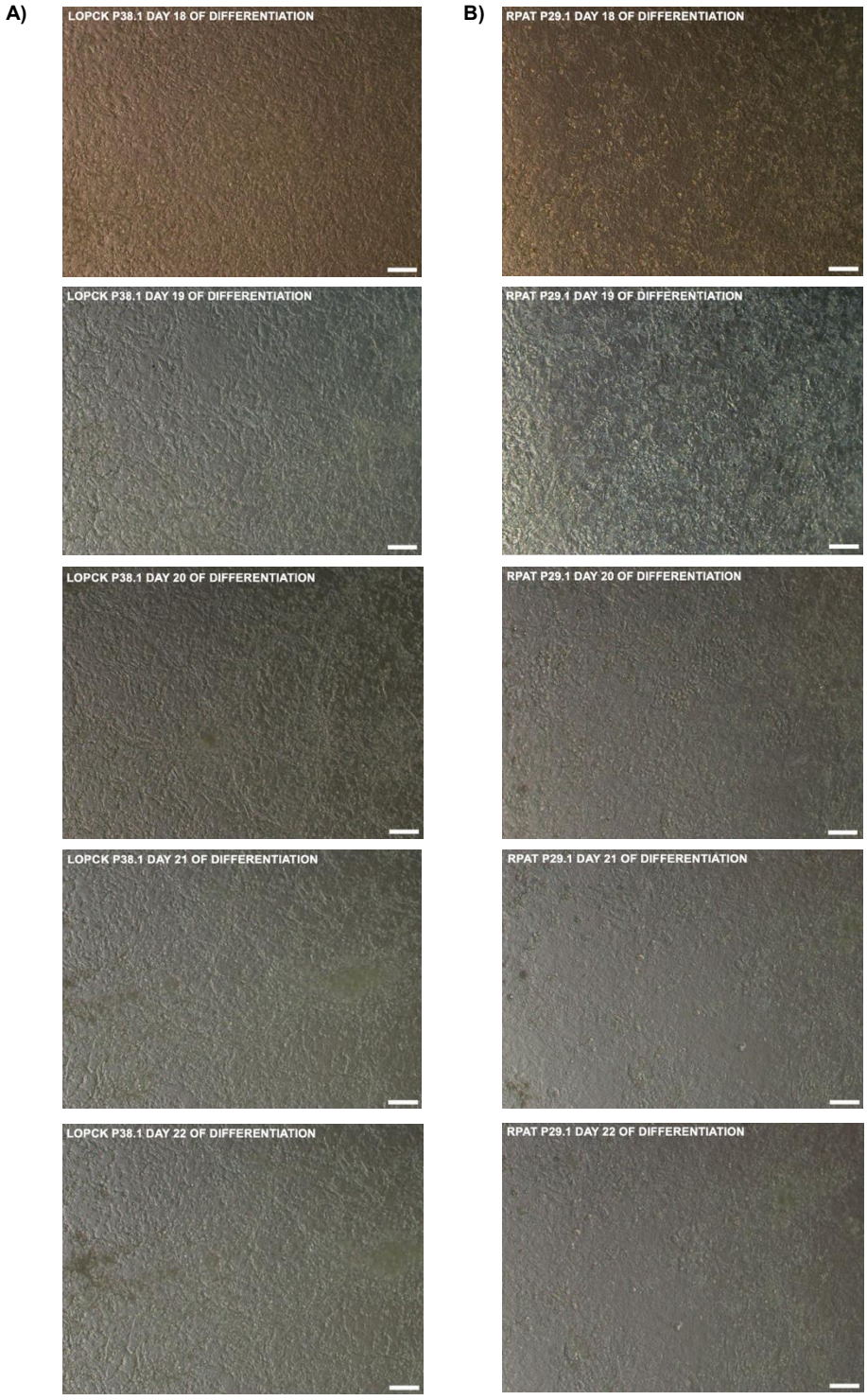


Figure 5-8. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 4 Expansion of Bronchial Lung Progenitors. A) LOPCK P38.1. B) REBL-PAT P29.1. Cells were cultured in RPMI media with BEGM media with FGF18. Scale bar: 100 µm. n = 3.

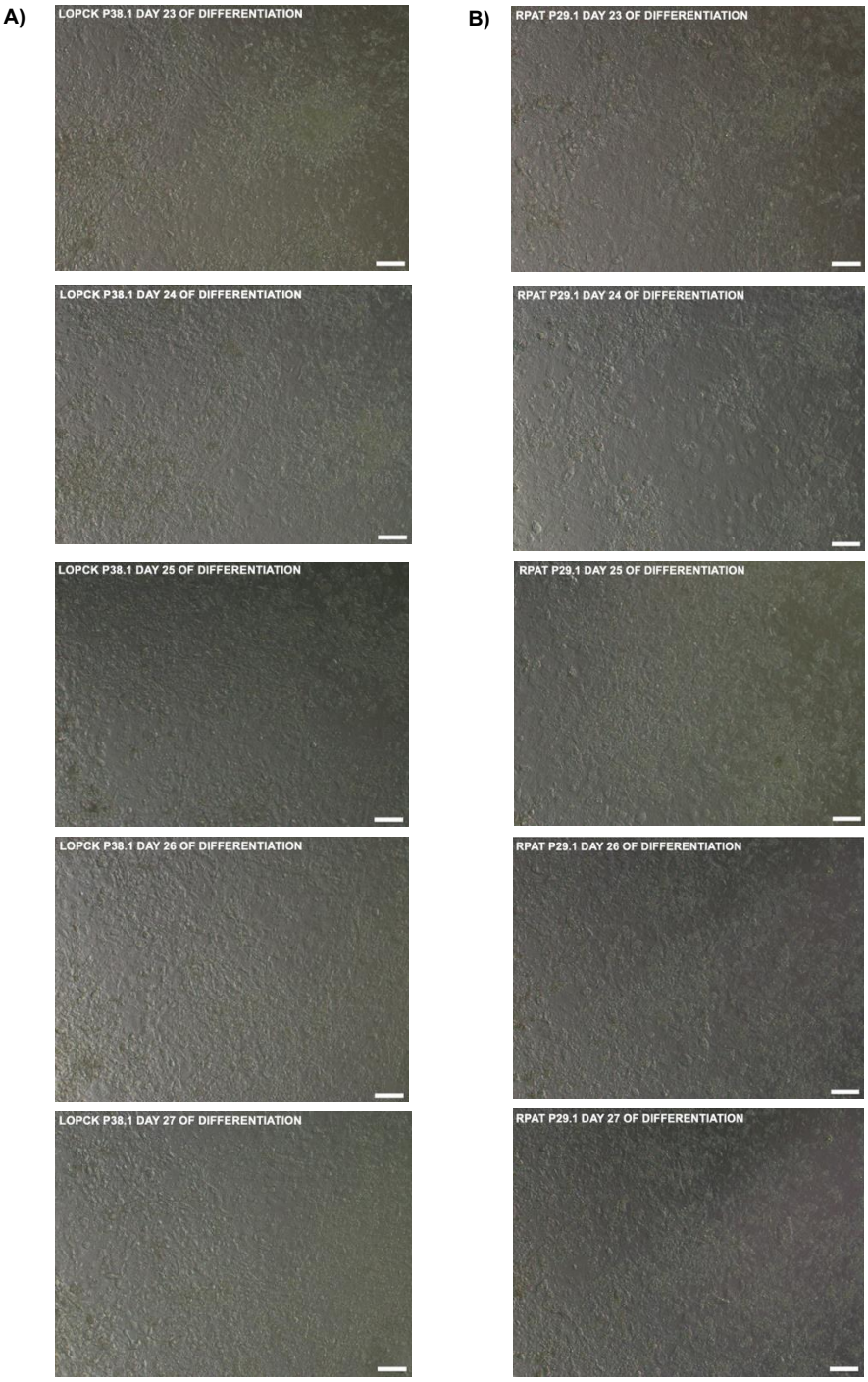


Figure 5-9. Images of LOPCK P38.1 and REBL-PAT P29.1 during Stage 4 Expansion of Bronchial Lung Progenitors. A) LOPCK P38.1. B) REBL-PAT P29.1. Cells were cultured in RPMI media with BEGM media with FGF18. Scale bar: 100 µm. n = 3.

5.4.2 D195307 HBECs Were Differentiated Into Mucociliary Pseudostratified Epithelium

To compare and contrast different models of the airway epithelium, HBEC-derived ALI model was established. D195307 Passage 2 HBECs were first cultured in submerged conditions until reaching confluency. Using PneumaCult^{TM}-ALI medium, HBECs were then sub-cultured onto porous-membrane transwell in ALI conditions. On day 4 mucus pits were detected (Figure 5-10). Swirls of mucus were observed indicating the development of mucus-secreting goblet cells. The ALI-HBECS cultures also exhibited ciliated cells which suggests the development of pseudostratified epithelium with ciliated, basal and goblet cells. To measure the barrier integrity, TEER readings were conducted on day 7, 14 and 21. The TEER measurements increased as the days in ALI culture increased (Figure 5-13). This increase was determined to be statistically significant across the days using repeated measures one-way ANOVA test (Day 7 vs Day 14 = $p < 0.0001$; Day 7 vs Day 21 = $p < 0.0001$; Day 14 vs Day 21 = $p < 0.0001$). Data was shown to be normally distributed using the Shapiro-Wilk test. This suggests the formation of tight junctions and increasing barrier integrity as the days in culture increased.

A representative image of haemotoxylin and eosin (H&E) and alcian blue and periodic acid-Schiff (PAS) staining of HBEC ALI at day 21 is shown in figure 5-14. The HBEC ALI transwell were fixed using formaldehyde and embedded in agarose prior to tissue processing. ALI sections were made and stained using haemotoxylin and eosin (H&E), and Alcian blue and periodic acid-Schiff (PAS) staining. The purplish-blue staining of the nuclei and pink staining of the extracellular matrix and cytoplasm was observed in sections stained with H&E staining (Figure 5-14 A-B). The light blue staining of the mucus was observed around the cells in sections stained with Alcian blue-PAS (Figure 5-14 C-D). This indicates that the HBECs were producing mucus.

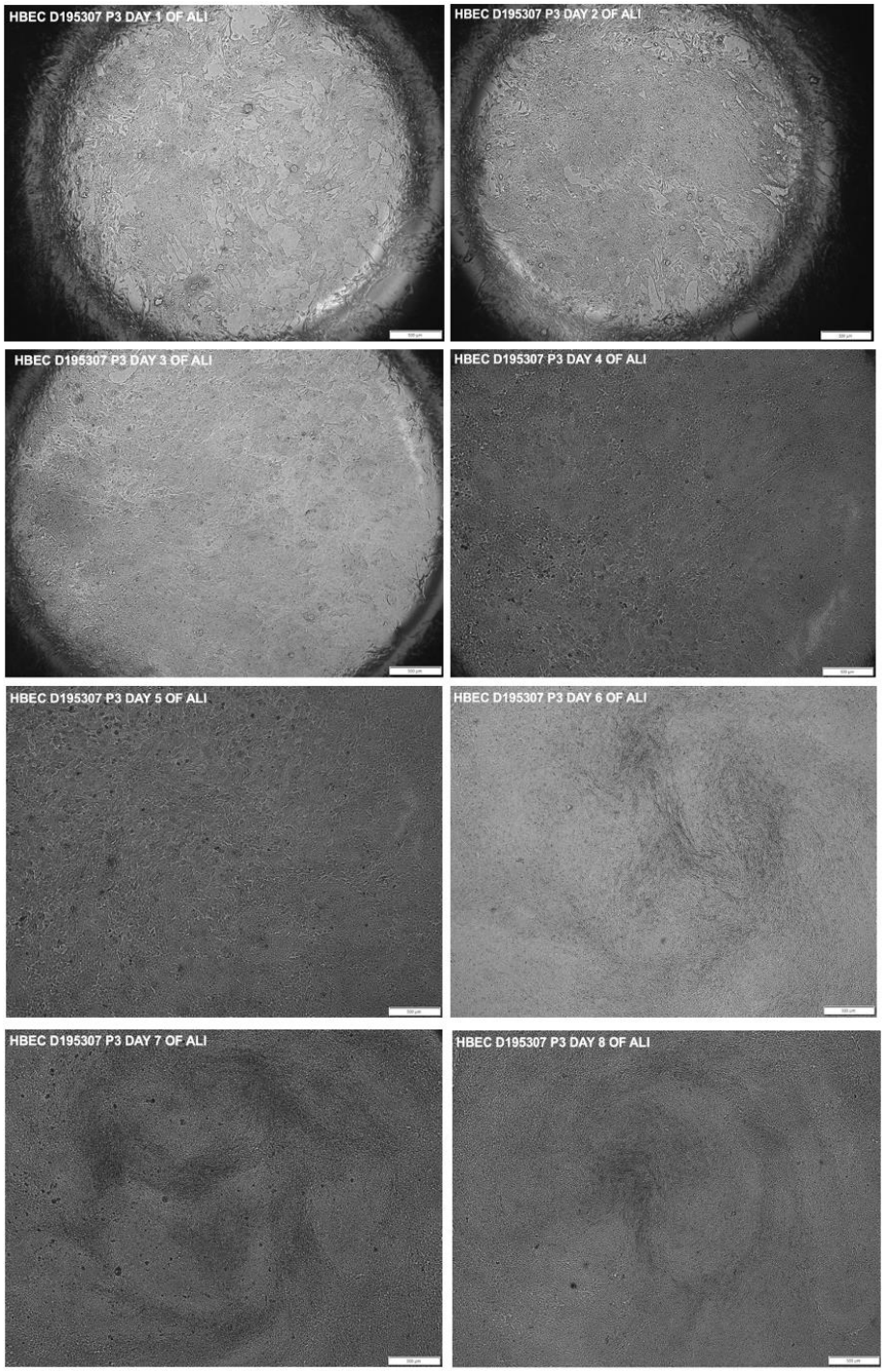


Figure 5-10. Images of D195307 Passage 3 HBECs cultured at ALI. HBECs were first cultured in PneumaCultTM-Ex Plus Medium until confluency. HBECs were then seeded into transwell and cultured using PneumaCult^{TM}-ALI medium. Scale bar: 500 µm. n = 3.

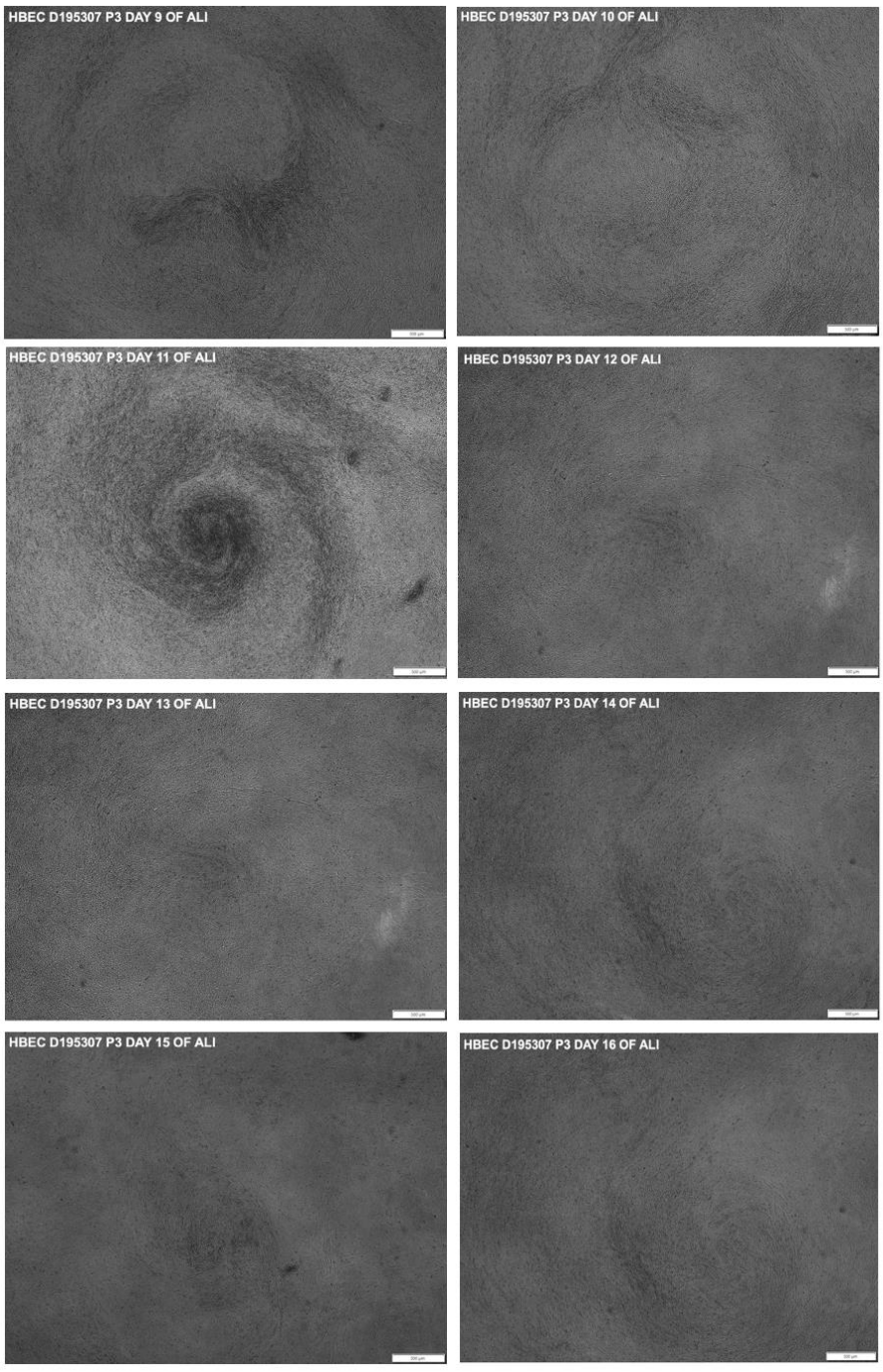


Figure 5-11. Images of D195307 Passage 3 HBECs cultured at ALI. HBECs were cultured in PneumaCultTM-Ex Plus Medium until confluency. HBECs were then seeded into transwell and cultured using PneumaCult^{TM}-ALI medium. Scale bar: 500 µm. n = 3.

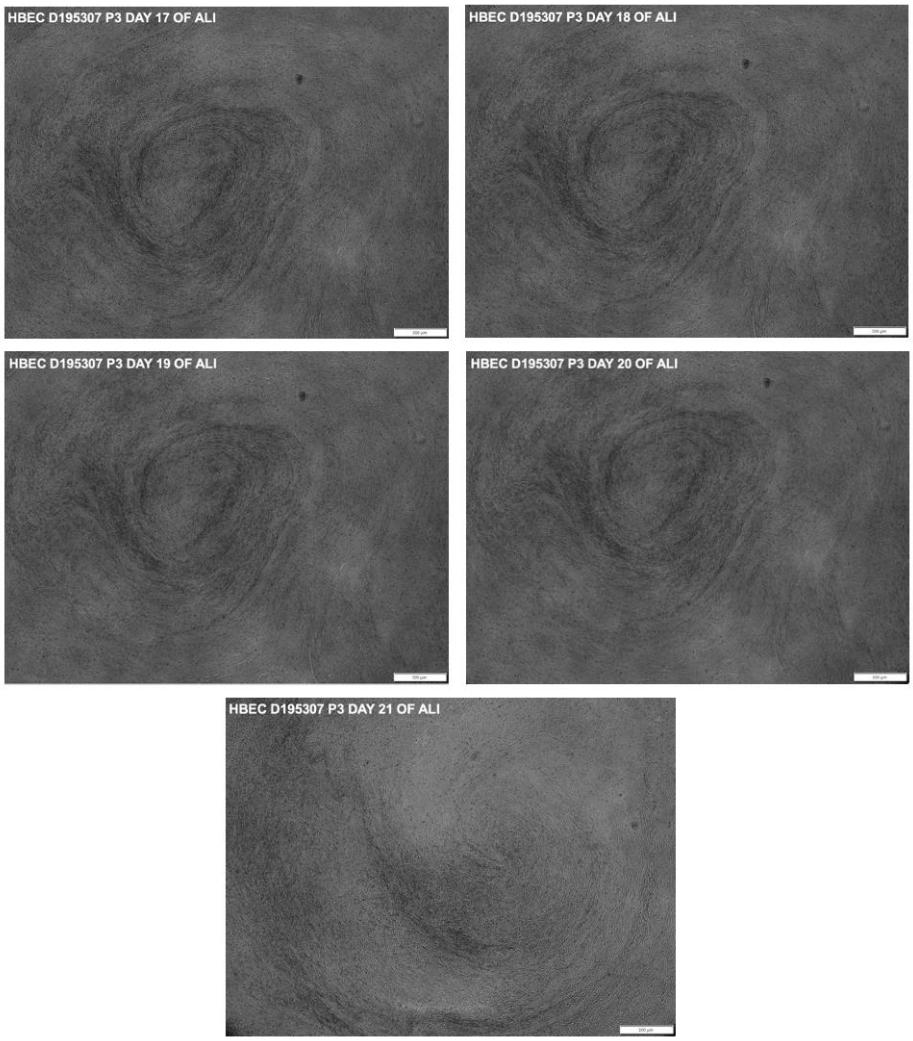


Figure 5-12. Images of D195307 Passage 3 HBECs cultured at ALI. HBECs were first cultured in PneumaCultTM-Ex Plus Medium until confluency. HBECs were then seeded into transwell and cultured using PneumaCult^{TM}-ALI medium. Scale bar: 500 µm. n = 3.

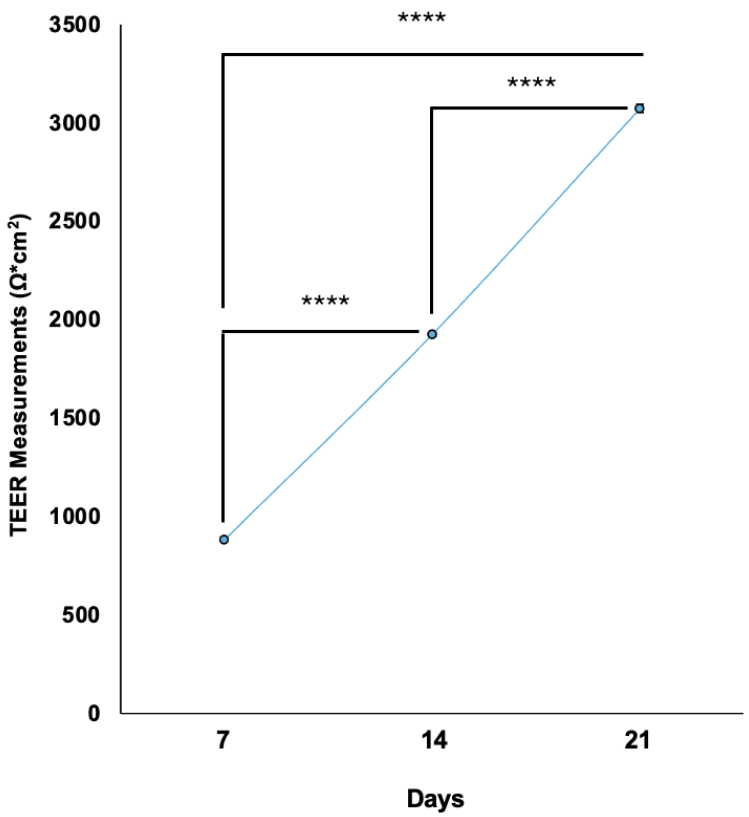


Figure 5-13. Increase in TEER measurements as the number of days in culture increased. TEER recordings were taken on day 7, 14 and 21 using an epithelial Voltohmmeter to assess the changes in barrier integrity. Repeated measures one-way ANOVA was used to compare values across the days. The increase in TEER readings was statistically significant across the group (**** = $p < 0.0001$). Shapiro-Wilk normality test was used to determine that the data was normally distributed. Error bars: mean ± SD. n = 3.

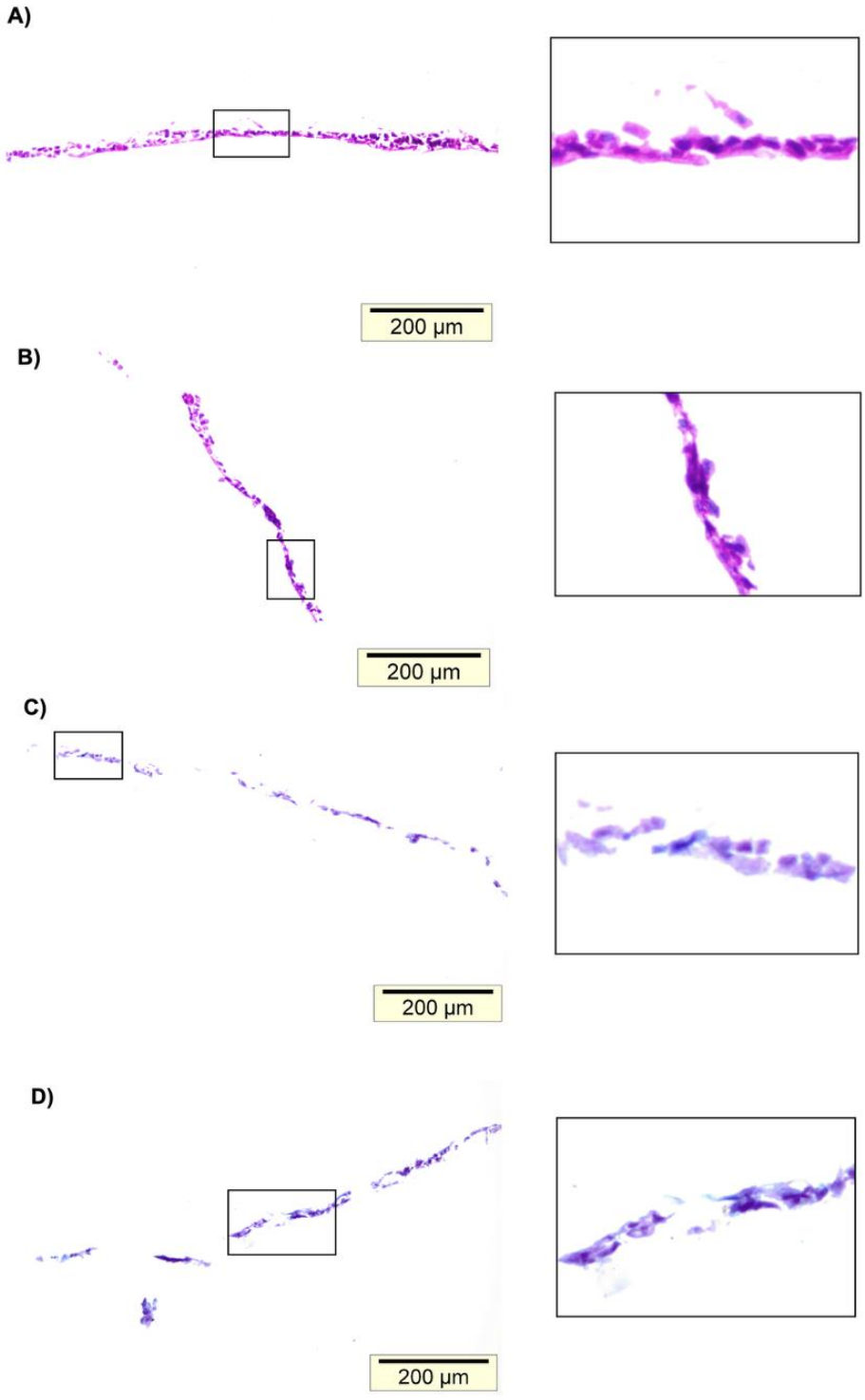


Figure 5-14. H&E staining and alcian-blue PAS staining of ALI-HBECs. A-B) H&E. **C-D:** Alcian-blue PAS staining. The insert was fixed on day 21, embedded and stained using H&E (A-B) and haemotoxylin, alcian blue, periodic acid and Schiff reagent (C-D) and viewed using a microscope at 10x lens. The boxes show the zoomed in images of the histology sections.

5.4.3 LOPCK and REBL-PAT iPSCs Were Unsuccessful in Forming a Mucociliary Pseudostratified Epithelium

After stage 4 of differentiation, the bronchial progenitor cells are cultured on ALI transwell using Lonza B-ALI™ Bronchial Air-Liquid Interface Medium BulletKit™ for 5 weeks according to the Wong *et al* protocol. The Lonza B-ALI™ media was used because the media was found to be the most reproducible for establishing ALI is iPSC cells by Wong *et al*[610]. However, the establishment of ALI was unsuccessful as the iPSC-derived bronchial lung progenitor cells failed to proliferate to form a monolayer (Figure 5-15B). The cultures were disintegrated and so the airway epithelium could not be formed. In contrast, ciliated cells can be identified in HBEC ALI cultures (Figure 5-15A). Mucus swirls and beating of the cilia could be observed under the microscope. This suggests the formation of a mature airway epithelium with ciliated and mucus-producing goblet cells. Therefore, HBEC-derived ALI culture was successfully differentiated into mature airway epithelium whereas iPSC-derived ALI cultures were unable to do so.

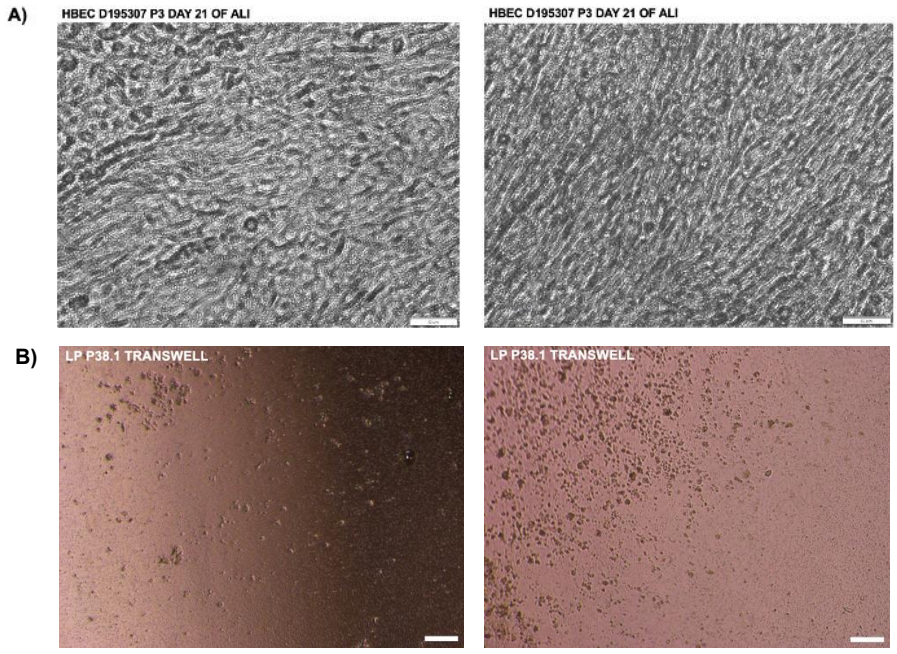


Figure 5-15. Comparison of HBEC-derived and iPSC-derived ALI cultures. A) HBEC-derived ALI culture where cilia can be identified. **B)** iPSC-derived ALI culture where cells failed to proliferate. n = 3.

5.4.4 Characterisation of LOPCK and REBL-PAT iPSCs at Different Stages of Differentiation

LOPCK and REBL-PAT cells were characterised at the end of each stage of differentiation where they were assayed for the following markers using RT-qPCR: FOXA2 (DE), SOX17 (DE), NKX2.1 (AFE), FOXJ1 (ciliated), KRT5 (basal), and MUC5AC (goblet). To compare the changes across the different stages, ΔCt value which was normalised to 18S was used. Due to non-detection of several markers at different stages of differentiation, Ct value for some markers could not be obtained and therefore fold gene expression across the different stages could not be calculated. This non-detection could be attributed to the amount of RNA used which was 50 ng/ml. This RNA quantity was due to the amount of iPSC cells that were available to be seeded into Transwell. Non-detects can indicate that the target was too low to be amplified therefore any low expression may be missed [640]. The ΔCt value of FOXA2 was lower in LOPCK Stage 1 (ΔCt = 15.79), than in Stage 2 (ΔCt = 16.40), Stage 3 (ΔCt = 18.14), and Stage 4 (ΔCt = 16.84) (Figure 5-16A). This was the case with REBL-PAT too where ΔCt was 12.03 for Stage 1 and 14.34 and 13.46 for Stage 2 and 3 respectively (Figure 5-16B). This suggests that the expression of FOXA2 is higher in Stage 1 compared to the other stages. However, one-way ANOVA showed no significance in FOXA2 expression across the different stages of LOPCK and REBL-PAT differentiation.

SOX17 was expressed during Stage 1 for both LOPCK (ΔCt = 19.81) and REBL-PAT (ΔCt = 16.56) suggesting formation of DE (Figure 5-17). SOX17 was not detected in any other stages of LOPCK and REBL-PAT differentiation (Figure 5-17). LOPCK showed expression of NKX2.1 during Stage 2 (ΔCt = 22.98) suggesting generation of AFE (Figure 5-18A). NKX2.1 was not detected in other stages of LOPCK and REBL-PAT differentiation (Figure 5-18). FOXJ1 was expressed during Stage 1 (ΔCt = 22.64), Stage 2 (ΔCt = 20.37) and Stage 3 (ΔCt = 19.30) of LOPCK differentiation (Figure 5-19A). In addition, FOXJ1 was also in Stage 1 (ΔCt = 18.24), Stage 2 (ΔCt = 17.99), and Stage 3 (ΔCt = 16.22) of REBL-PAT differentiation (Figure 5-19B). Since FOXJ1 is a marker of ciliated cells, this suggests that spontaneous differentiation may have possibly occurred. However, the changes in FOXJ1 expression across the

different stages was non-significant as determined by one-way ANOVA. KRT5 and MUC5AC were not detected at any stage of LOPCK and REBL-PAT differentiation. FOXJ1 (ΔCt = 11.47), KRT5 (ΔCt = 12.19), and MUC5AC (ΔCt = 11.46) was detected in ALI-HBECs suggesting differentiation of HBECs into a pseudostratified airway epithelium complete with ciliated cells, basal and mucus-producing goblet cells.

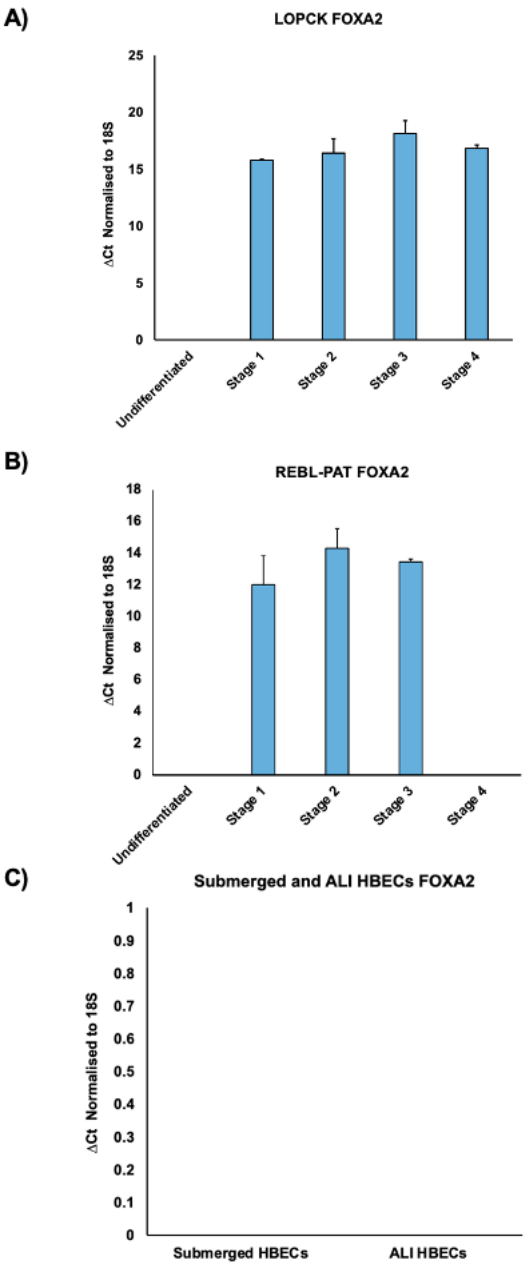


Figure 5-16. Stage 1 cells showed lower ΔCT value normalised to 18S of FOXA2 in LOPCK and REBL-PAT cells demonstrating formation of DE. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs. One-way ANOVA with Tukey's multiple comparison test showed no significance between ΔCT value at different stages of LOPCK and REBL-PAT differentiation. Error bars: mean ± SD. n = 3.

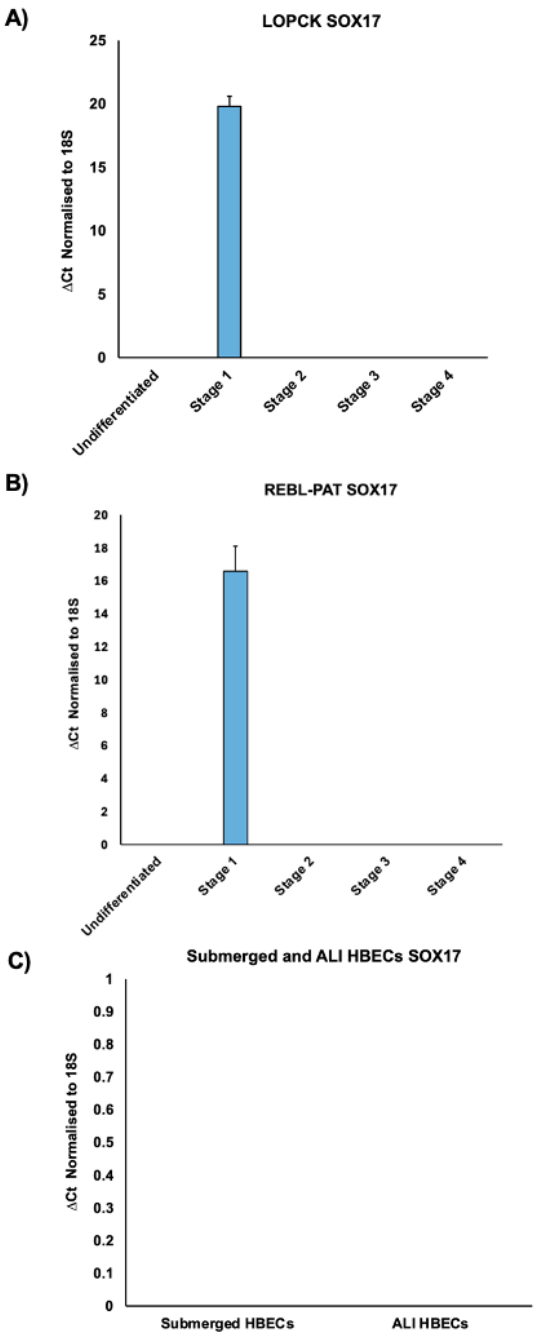


Figure 5-17. Stage 1 cells showed ΔCT value normalised to 18S of SOX17 in LOPCK and REBL-PAT cells demonstrating formation of DE. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs. Blank bars represents non-detection of the marker. Error bars: mean ± SD. n = 3

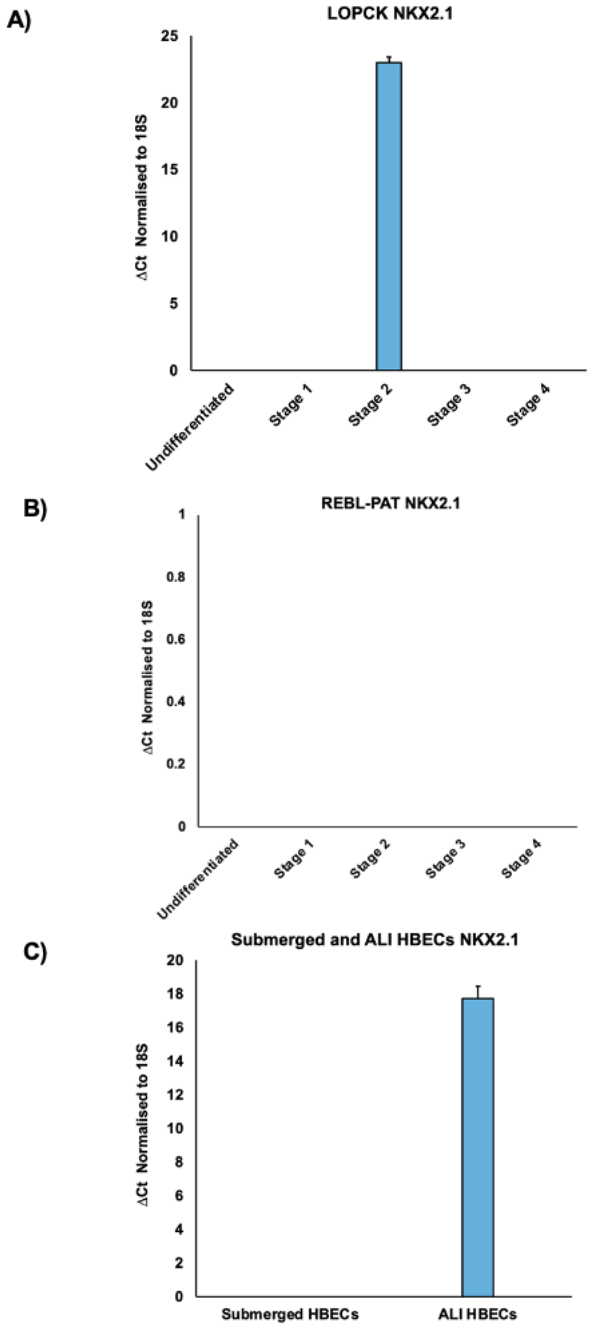


Figure 5-18. Stage 1 cells showed ΔCT value normalised to 18S of NKX2.1 in LOPCK cells demonstrating formation of AFE. A) LOPCK. **B)REBL-PAT**. **C)** Submerged and ALI-HBECs. Blank bars represents non-detection of the marker. Error bars: mean ± SD. n = 3

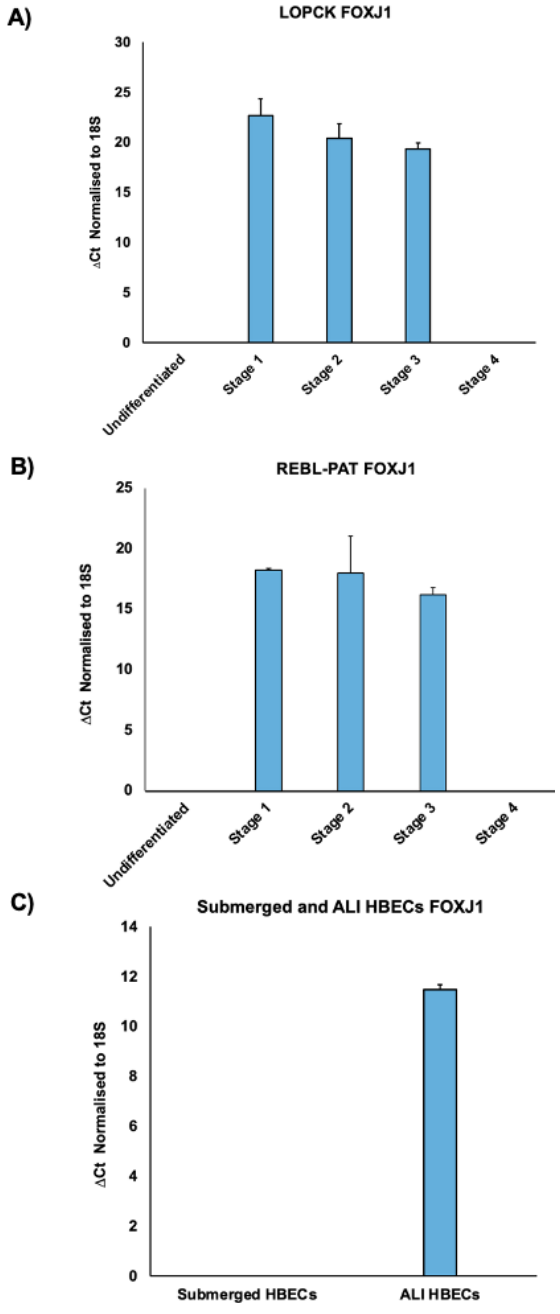


Figure 5-19. ΔCT values normalised to 18S of FOXJ1 in iPSCs, submerged and ALI HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs. Two-way ANOVA test with Tukey's multiple comparison test showed no significance for the marker across different stages of LOPCK and REBL-PAT differentiation. Error bars: mean ± SD. n = 3

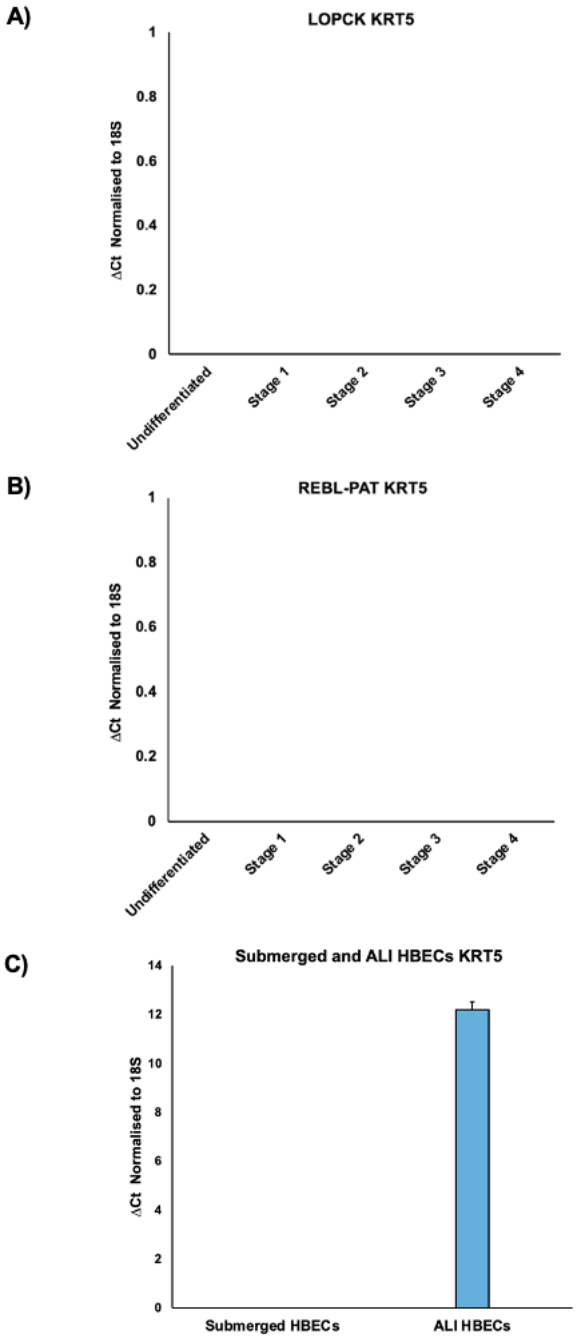


Figure 5-20. ΔCT values normalised to 18S of KRT5 at different stages of iPSC differentiation, submerged and ALI HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs. Blank bars represents non-detection of the marker. Error bars: mean ± SD. Error bars: mean ± SD. n = 3

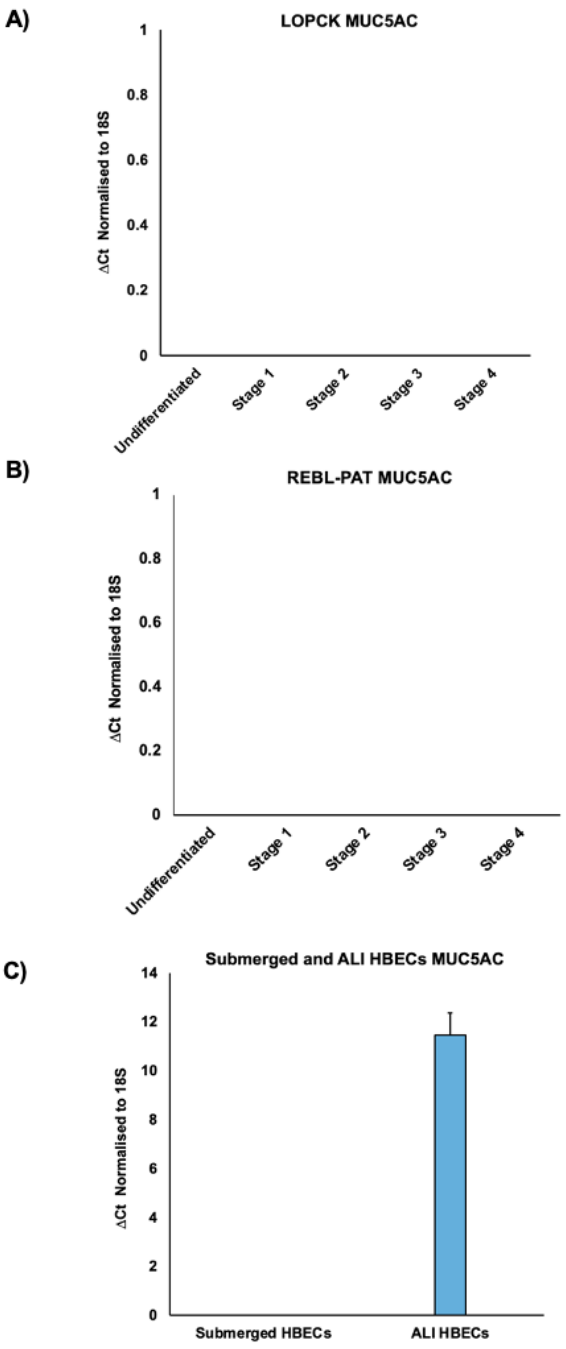


Figure 5-21. ΔCT values normalised to 18S of MUC5AC at different stages of iPSC differentiation, submerged and ALI HBEC cultures. A) LOPCK. **B)REBL-PAT**. **C)** Submerged and ALI-HBECs**.** Blank bars represents non-detection of the marker. Error bars: mean ± SD. n = 3

5.4.5 Investigating How Permissive iPSCs and HBECs Are To RSV-A2 At Various Stages Of Differentiation

Host-virus interaction at each stage of differentiation was investigated to see how permissive the cells were to RSV-A2. Viral load was measured using RT-qPCR. The changes in the expression levels of RSV receptor CX3CR1, interferons and restriction factors OAS1 were also studied across various stages in the presence and absence of RSV. LOPCK, REBL-PAT iPSCs and D195307 HBECs were infected with RSV A2 at an MOI of 10. The virus was removed after 1 hour incubation and replaced with fresh media. After 48 hours, cells were lysed and RNA extracted to measure gene expression levels using RT-qPCR. ΔCT values for RSV receptor CX3CR1 was lower in RSV-infected undifferentiated, Stage 1 and Stage 3 cells compared to non-infected undifferentiated, Stage 1 and Stage 3 LOPCK cells (Figure 5-22A). However, analysis using two-way ANOVA showed no significant differences between the non-infected and infected cells across the different stages. Whereas the ΔCT values were lower for CX3CR1 in RSV-infected submerged (ΔCT = 11.66) and ALI-HBECs (ΔCT = 10.11) compared to non-infected submerged (ΔCT = 13.69) and ALI-HBECs (ΔCT = 12.40) (Figure 5-22C). This suggests that RSV infection can lead to higher expression of CX3CR1 in submerged and ALI-HBECs. This relationship was found to be statistically significant ($p \leq 0.05$).

Upon RSV infection, IFNβ was expressed in submerged (ΔCT = 14.21) and ALI-HBECs (ΔCT = 11.29) (Figure 5-23C). IFNλ was also expressed in RSV-infected submerged HBECs (ΔCT = 16.72) (Figure 5-24C). Statistical test could not be carried out for IFNβ and IFNλ in submerged and ALI-HBECs due to non-detection of the markers in the non-infected submerged and ALI-HBECs. The ΔCT values were lower for OAS1 in RSV-infected submerged (ΔCT = 9.71) and ALI-HBECs (ΔCT = 9.56) compared to non-infected submerged (ΔCT = 11.54) and ALI-HBECs (ΔCT = 11.38) (Figure 5-25C). This suggests that expression of OAS1 increased during RSV infection in submerged and ALI-HBECs. This relationship was found to be statistically significant ($p \leq 0.05$). RSV-A2 was not detected in any stages of LOPCK, REBL-PAT differentiation, and in submerged or ALI-HBECs (Figure 5-26).

However, the detection of interferon response and OAS1 following RSV infection in submerged and ALI HBEC suggest that the cells may have been infected but the viral load of RSV-A2 is too low for detection. The expression of CX3CR1, IFNβ and OAS1 varied between non-infected and infected LOPCK and REBL-PAT cells. Analysis using two-way ANOVA for comparison across groups showed no significance in LOPCK and REBL-PAT cells. This suggests that the change in the expression of markers is likely not due to RSV infection.

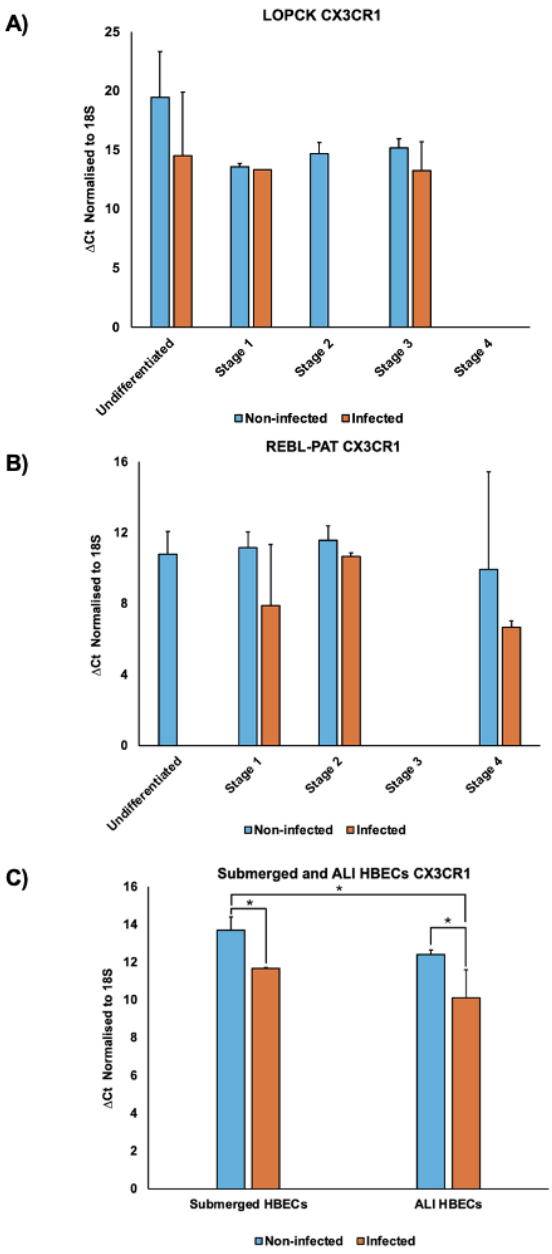


Figure 5-22. ΔCT values normalised to 18S of CX3CR1 in non-infected and RSV infected iPSC and HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs**.** Cells were infected with RSV MOI of 10 for 1 hour and lysed after 48 hours for RT-qPCR. Two-way ANOVA with Tukey's multiple comparison test was used to compare values across group. * = $p \leq 0.05$. Error bars: mean ± SD. n = 3

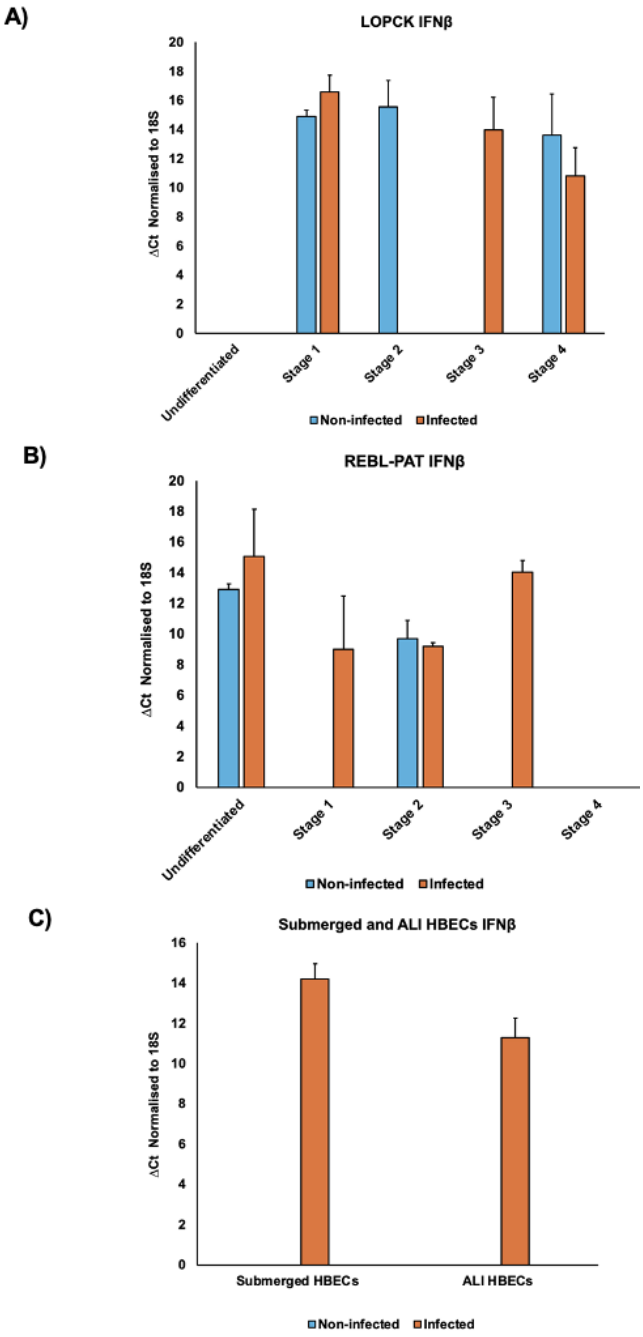


Figure 5-23. ΔCT values normalised to 18S of IFNβ in non-infected and RSV infected iPSC and HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs**.** Cells were infected with RSV MOI of 10 for 1 hour and lysed after 48 hours for RT-qPCR. Two-way ANOVA test with Tukey's multiple comparison test showed no significance in values across group. Error bars: mean ± SD. n = 3

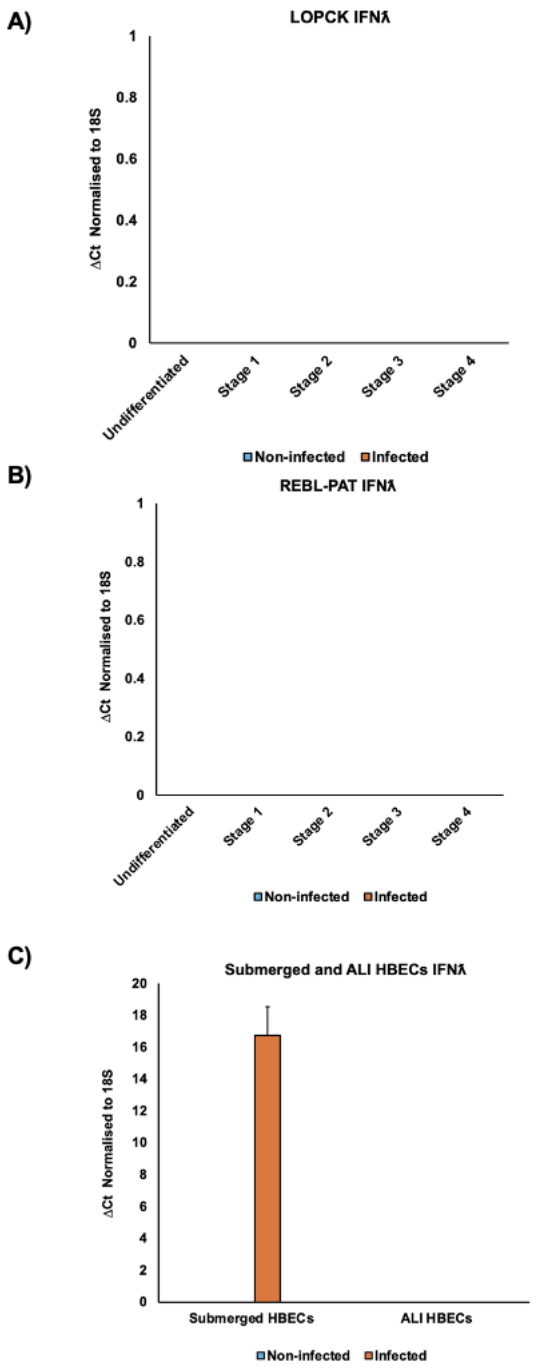


Figure 5-24. ΔCT values normalised to 18S of IFNλ in non-infected and RSV infected iPSC and HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs**.** Cells were infected with RSV MOI of 10 for 1 hour. After 48 hours, cells were lysed for RT-qPCR. Blank bars represents non-detection of the marker. n = 3

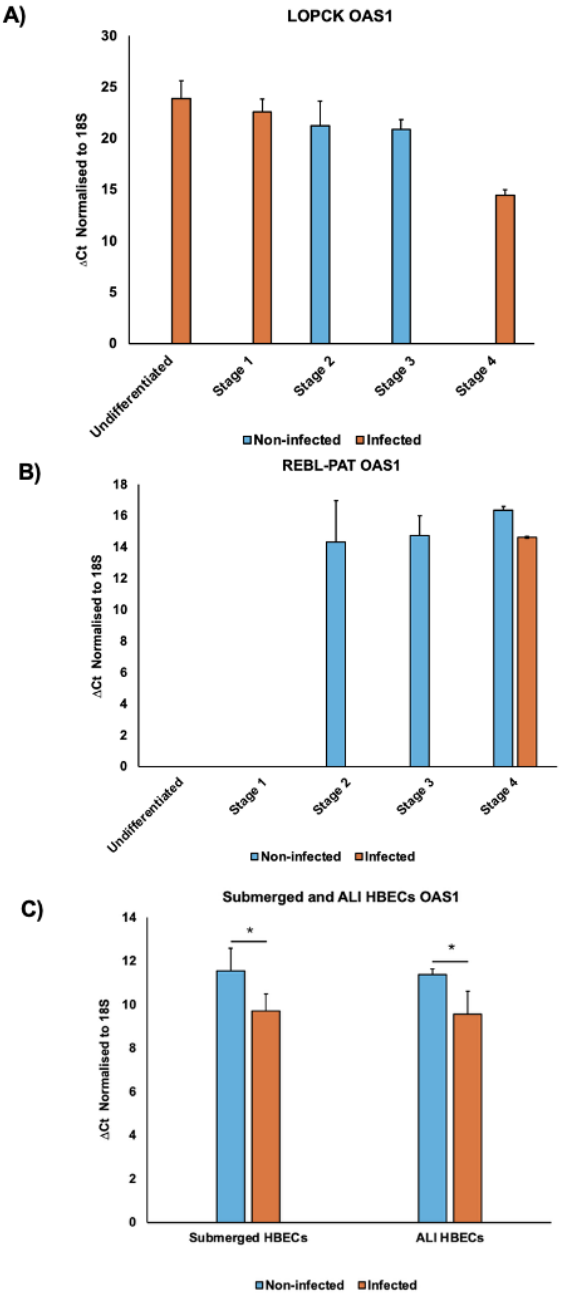


Figure 5-25. ΔCT values normalised to 18S of OAS1 in non-infected and RSV infected iPSC and HBEC cultures. A) LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs**.** Cells were infected with RSV MOI of 10 for 1 hour and lysed after 48 hours for RT-qPCR. Two-way ANOVA test with Tukey's multiple comparison test was used to compare values across group. * = $p \leq 0.05$. Error bars: mean ± SD. n = 3

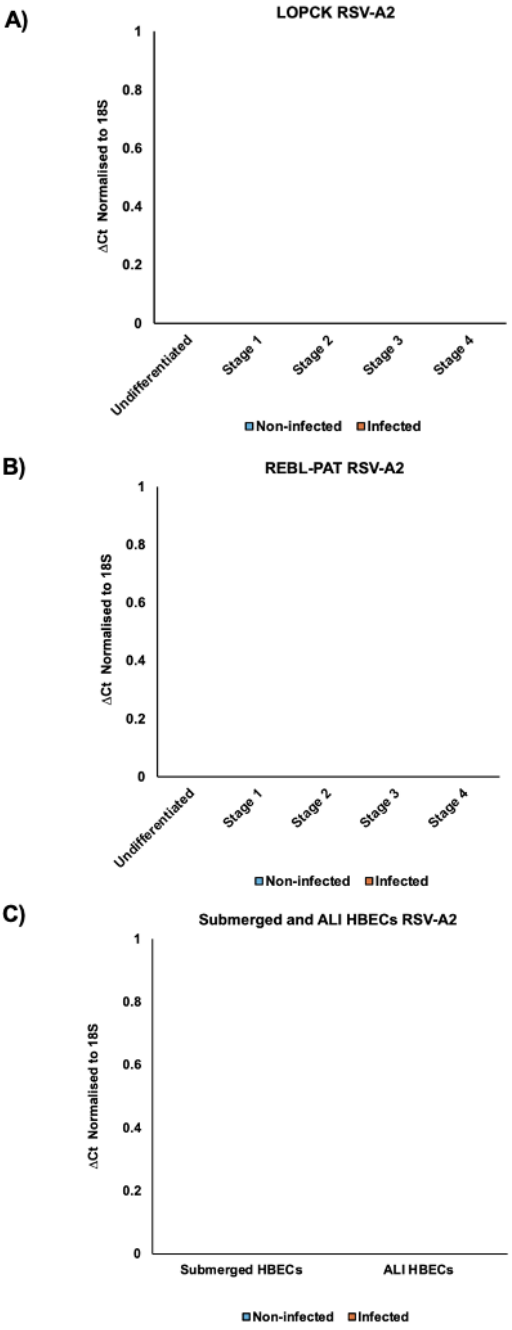


Figure 5-26. ΔCT values normalised to 18S of RSV-A2 in non-infected and RSV infected iPSC and HBEC cultures. **A)** LOPCK. **B)** REBL-PAT. **C)** Submerged and ALI-HBECs**.** Cells were infected with RSV MOI for 1 hour and lysed after 48 hours for RT-qPCR. Blank bars represents non-detection of the marker. n = 3.

To further investigate how permissive iPSCs are at different stages of differentiation, flow cytometry was utilised in preliminary experiments to the number of RSV-infected cells. LOPCK cells were infected with RSV-A2 at an MOI of 10. The RSV virus was tagged with green fluorescent protein (GFP). Following incubation with RSV A2 for 1 hour, virus was removed, and fresh media was added. The cells were detached and fixed in PFA after 48 hours. Flow cytometry was used to measure the percentage of infection between infected and non-infected cells. The mean percentage of GFP cells were similar between non-infected and infected undifferentiated and stage 1 DE cells. The mean percentage of GFP-positive cells was higher in RSV-infected stage 2 cells (0.03%) in comparison to non-infected stage 2 cells (9.23%) (Figure 5-29). This increase in similarity was echoed in RSV-infected stage 3 cells. There was a bigger increase in GFP-positive cells in RSV-infected stage 4 cells (44.51%) compared to non-infected stage 4 cells (2.75%) (Figure 5-29) These results suggest that LOPCK cells become more permissive as the stages of differentiation progresses. A two-way ANOVA test with Tukey's multiple comparison test was used to compare values across groups (Table 5-16). The changes in the mean percentage of GFP cells between each stage of differentiation and their infected counterpart was found to be statistically significant.

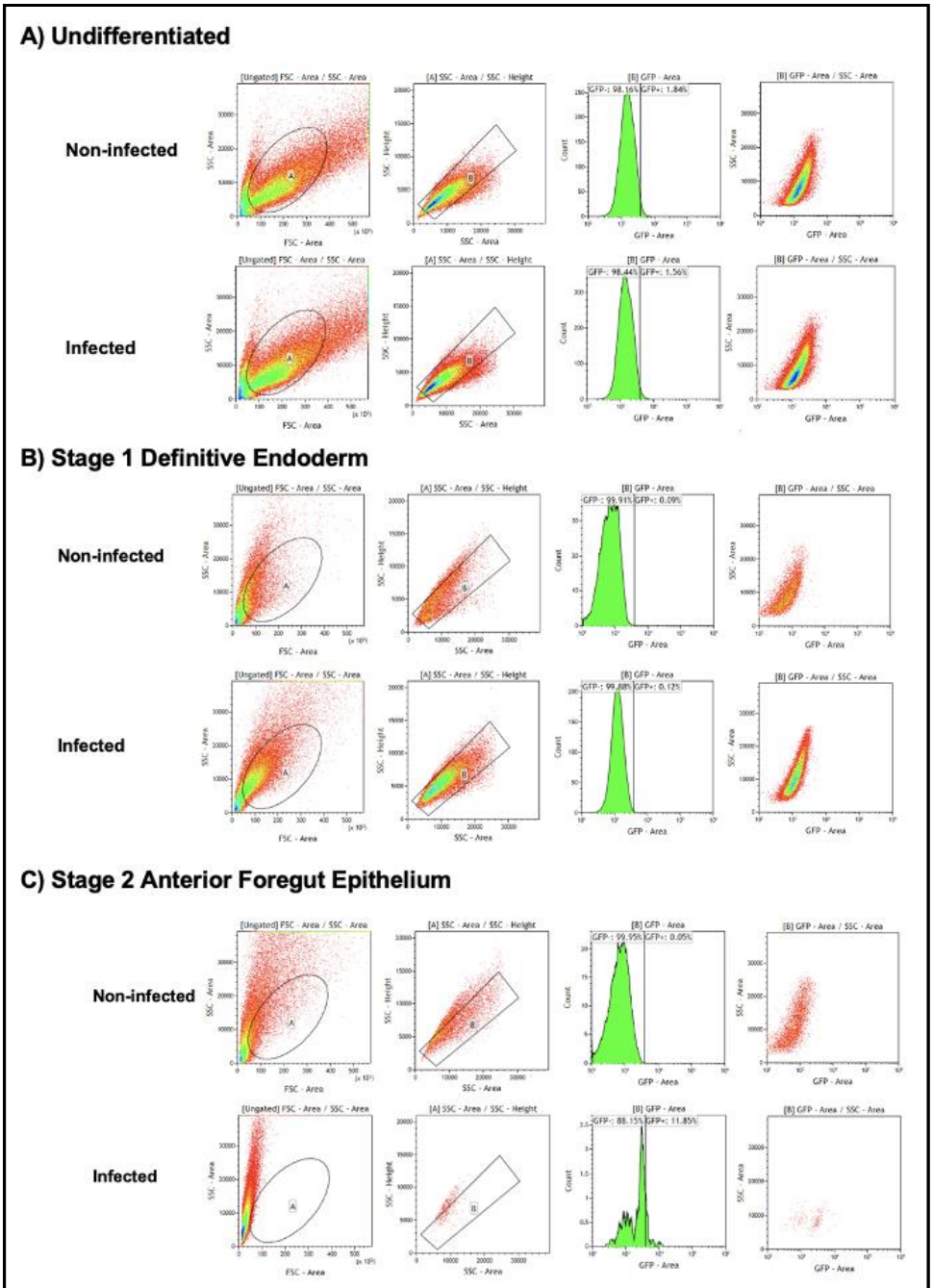


Figure 5-27. Flow cytometry of non-infected and RSV-GFP infected LOPCK cells at undifferentiated stage and stages 1 and 2 of differentiation. A) Undifferentiated. **B)** Stage 1 Definitive Endoderm. **C)** Stage 2 Anterior Foregut Epithelium. Cells were infected with RSV-GFP at MOI of 10 for 1 hour. After 48 hours, cells were fixed in PFA, and flow cytometry was conducted. n = 3.

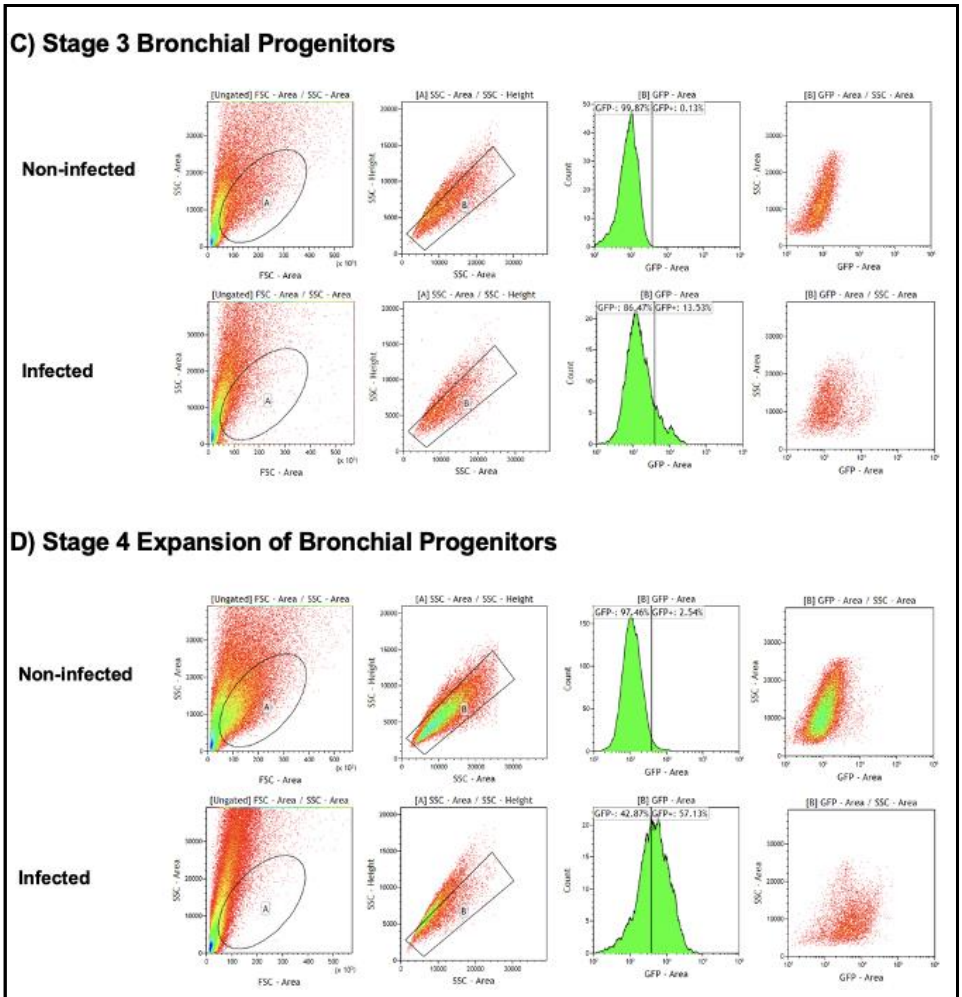


Figure 5-28. Flow cytometry of non-infected and RSV-GFP infected LOPCK cells at stages 3 and 4 of differentiation. A) Stage 3 Bronchial Progenitors. **B)** Stage 4 Expansion of Bronchial Progenitors. Cells were infected with RSV-GFP at MOI of 10 for 1 hour. After 48 hours, cells were fixed in PFA, and flow cytometry was conducted. n = 3.

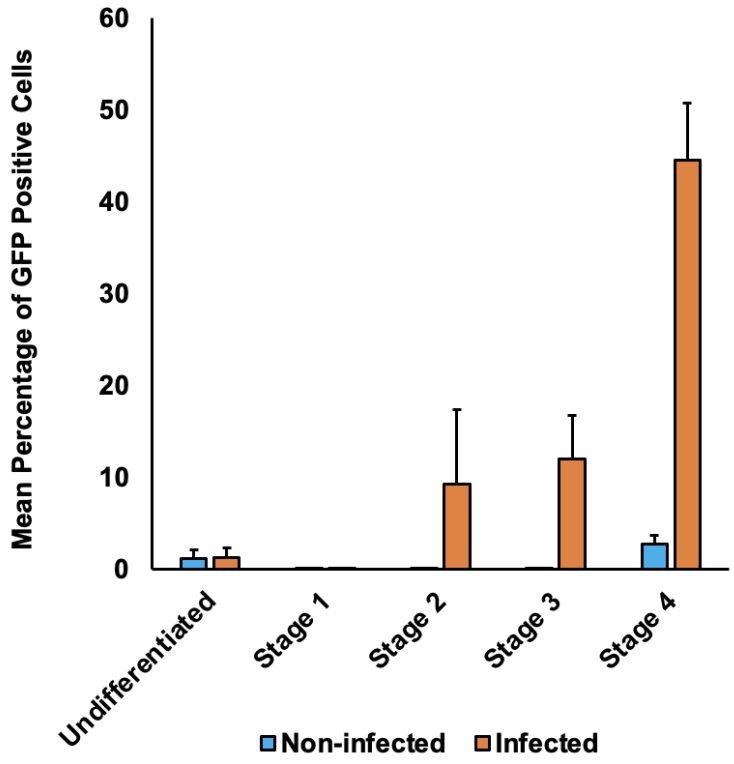


Figure 5-29. Percentage of GFP-positive cells within non-infected and RSV-GFP infected LOPCK cells. A) Non-infected LOPCK cells. **B)** Infected LOPCK cells. Cells were infected with RSV-GFP at MOI of 10 for 1 hour. After 48 hours, cells were fixed in PFA, and flow cytometry was conducted. Two-way ANOVA test with Tukey's multiple comparison test was used to compare values across group which is shown in table 16. * = $p \leq 0.05$. Error bars: mean ± SD. n = 3.

Table 5-16. Table of multiple comparison between different groups which showed significance when LOPCK cells were infected with RSV-GFP. Two-way ANOVA test with Tukey's multiple comparison test was used to compare values across groups.

	Adjusted p-value
Undifferentiated:Non-infected vs. Undifferentiated:Infected	0.0272
Undifferentiated:Non-infected vs. Stage 4:Non-infected	0.009
Undifferentiated:Non-infected vs. Stage 4:Infected	0.0003
Undifferentiated:Infected vs. Stage 4:Infected	0.009
Stage 1:Non-infected vs. Stage 1:Infected	0.0272
Stage 1:Non-infected vs. Stage 4:Non-infected	0.0054
Stage 1:Non-infected vs. Stage 4:Infected	0.0002
Stage 1:Infected vs. Stage 4:Infected	0.0054
Stage 2:Non-infected vs. Stage 2:Infected	0.0272
Stage 2:Non-infected vs. Stage 4:Non-infected	0.0397
Stage 2:Non-infected vs. Stage 4:Infected	0.0013
Stage 2:Infected vs. Stage 4:Infected	0.0397
Stage 3:Non-infected vs. Stage 3:Infected	0.0272
Stage 3:Non-infected vs. Stage 4:Infected	0.0023
Stage 4:Non-infected vs. Stage 4:Infected	0.0272

5.4.6 Comparison Between The Way Human Epithelial Cells Respond To RV, RSV and SARS-CoV-2 Infections in ALI-HBECs Model of the Airway Epithelium

Using publicly available data, the similarities and differences between the way human epithelial cells respond to RV, RSV and SARA-CoV-2 infections in ALI-HBECs were examined. A study by Bai *et al* was used to obtained differently expressed genes in response to 24-hour RV infection of ALI-HBECs[641]. For RSV, a study by Anderson *et al* was used to attain differently expressed genes in response to 4-day RSV infection of ALI-HBECs and Vaid *et al* for 4-day SARS-CoV-2 infection of ALI-HBECs[642,643]. 210, 2 and 2 significantly upregulated genes were found to overlap between RV/RSV, RV/SARS-CoV-2 and RSV/SARS-CoV-2 respectively (Figure 5-30A). 10 upregulated genes were found to be common between all three viruses. Majority of the overlapping genes between the three viruses encodes for proteins that are involved in immune regulation or defences against viral infection. These include the chemokines CCL5, C-X-C motif chemokine ligand 10 (CXCL10) and C-X-C motif chemokine 11 (CXCL11), mediator of anti-viral immunity oligoadenylate synthetase-Like (OASL)[644], sensor of dsDNA and RNA Z-DNA binding protein 1 (ZBP1)[645], regulators of cytokine signalling the suppressor of cytokine signaling 1 (SOCS1)[646], and a key player of anti-viral host defence apolipoprotein B MRNA editing enzyme catalytic subunit 3A (APOBEC3A)[647].

Zinc finger protein 36 homolog (ZFP36) which is involved in regulation of T-cell activation, was found to be unique for RV. Interferon alpha inducible protein 27 (IFI27) and NLR family pyrin domain containing 10 (NLRP10), which is involved in apoptosis and immune regulation, were unique for RSV and SARS-CoV-2 respectively[648]. For downregulated genes, 16, 362 and 17 genes were RV/RSV, RV/SARS-CoV-2 and RSV/SARS-CoV-2 respectively (Figure 5-30B). Whereas 22 significantly downregulated genes were common for all three viruses. These include tubulin tyrosine ligase Like 7 (TTLL7) which is involved in β-tubulin polyglutamylation[649], DSTYK which encodes a dual serine/threonine and tyrosine kinase[650], and solute carrier family 23 member 2

(SLC23A2) which is implicated in transportation of vitamin C[651]. For downregulated genes unique to the viruses, negative elongation factor complex member C/D NELFCD/TH1L which regulates pausing and transcription elongation of RNA polymerase II was unique to RV[652]. Asporin (ASPN), which encodes for extracellular matrix protein and is implicated in osteoarthritis and cancer was found to be unique for RSV[653]. Solute carrier family 34 member (SLC34A) transporter family of proteins-encoding gene SLC34A3 was unique for SARS-CoV-2[654].

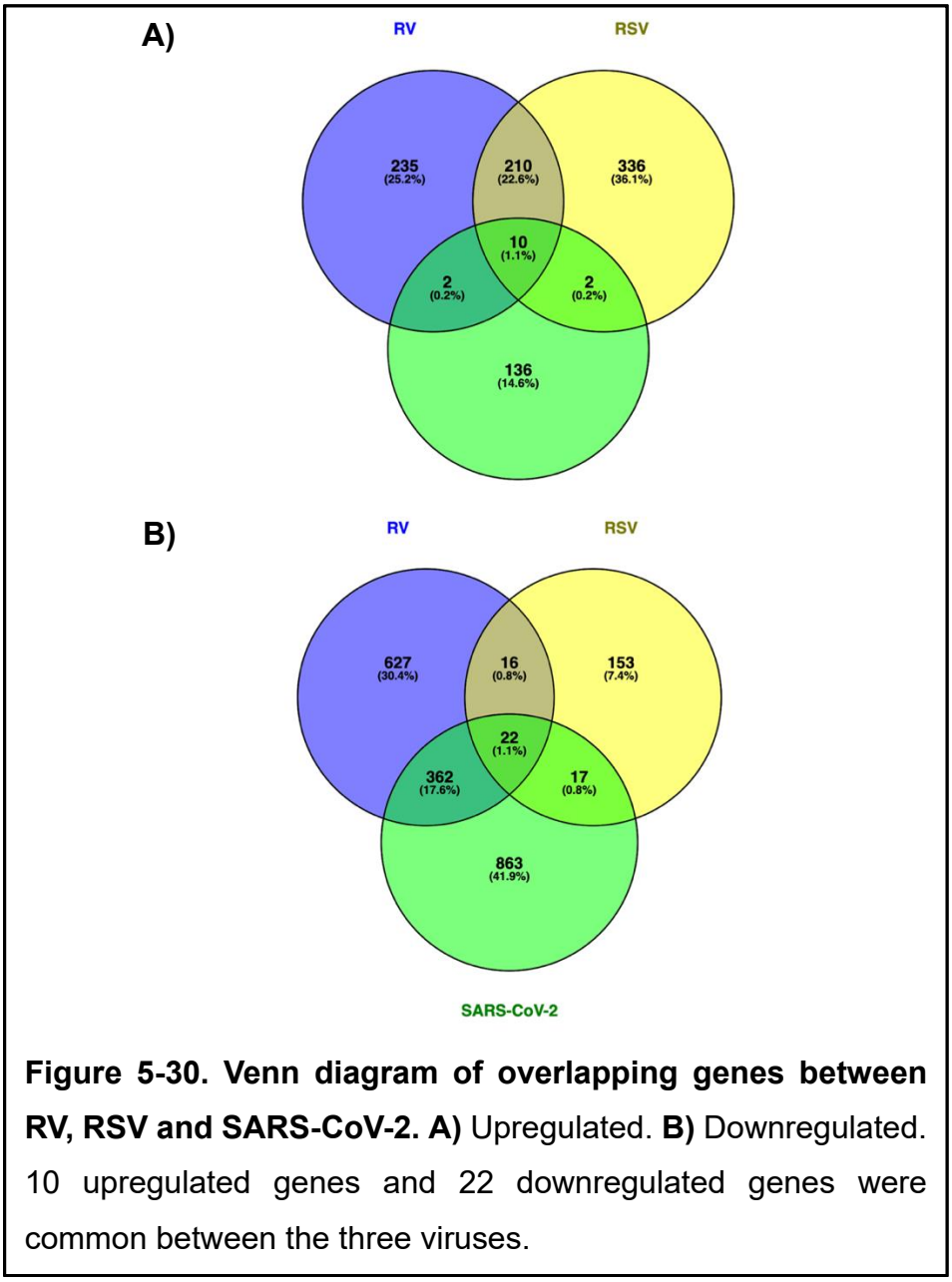


Figure 5-30. Venn diagram of overlapping genes between RV, RSV and SARS-CoV-2. A) Upregulated. **B)** Downregulated. 10 upregulated genes and 22 downregulated genes were common between the three viruses.

The aforementioned studies employed Gene Ontology (GO) Biological Process to conduct pathway analysis of upregulated genes. 253, 58 and 137 pathways were common for RV/RSV, RV/SARS-CoV-2 and RSV/SARS-CoV-2 respectively (Figure 5-31A). 187 pathways enriched for upregulated genes were found to be common between the three viruses. Pathways involving immune system such as regulation of innate immune response, positive regulation of defence response, positive regulation of cytokine production, regulation of chemokine production, regulation of type I interferon production, and positive regulation of lymphocyte activation were found to overlap between RV/RSV/SARS-CoV-2. Pathways implicating viral infection were also common for the three viruses. These include positive regulation of defence response to virus by host, regulation of defence response to virus, defence response to virus, regulation of viral entry into host cell, and cellular response to virus. Regulation of intracellular signal transductions was one of the pathways enriched for upregulated genes that was unique to RV. Regulation of cellular catabolic process was unique to RSV whereas calcium ion transport was unique to SARS-CoV-2.

Pathway analysis of downregulated genes revealed 84 common pathways for RSV/SARS-CoV-2 (Figure 5-31B). Moreover, 26 pathways were found to overlap between the three viruses. Pathways involving cilia such as cilium assembly, cilium organization, and protein localization to cilium were enriched for downregulated genes for all three viruses. Furthermore, pathways implicating the cytoskeleton such as microtubule-based movement, microtubule-based process, cytoskeleton-dependent intracellular transport, microtubule-based transport, microtubule cytoskeleton organization, and cytoskeleton organization were also common between the three viruses. There was zero pathways enriched for downregulated genes between RV/RSV and RV/SARS-CoV-2. Furthermore, there was zero pathways enriched for downregulated genes that were unique for RSV and SARS-CoV-2. Pathways unique to RV included cellular component organization and cellular component biogenesis.

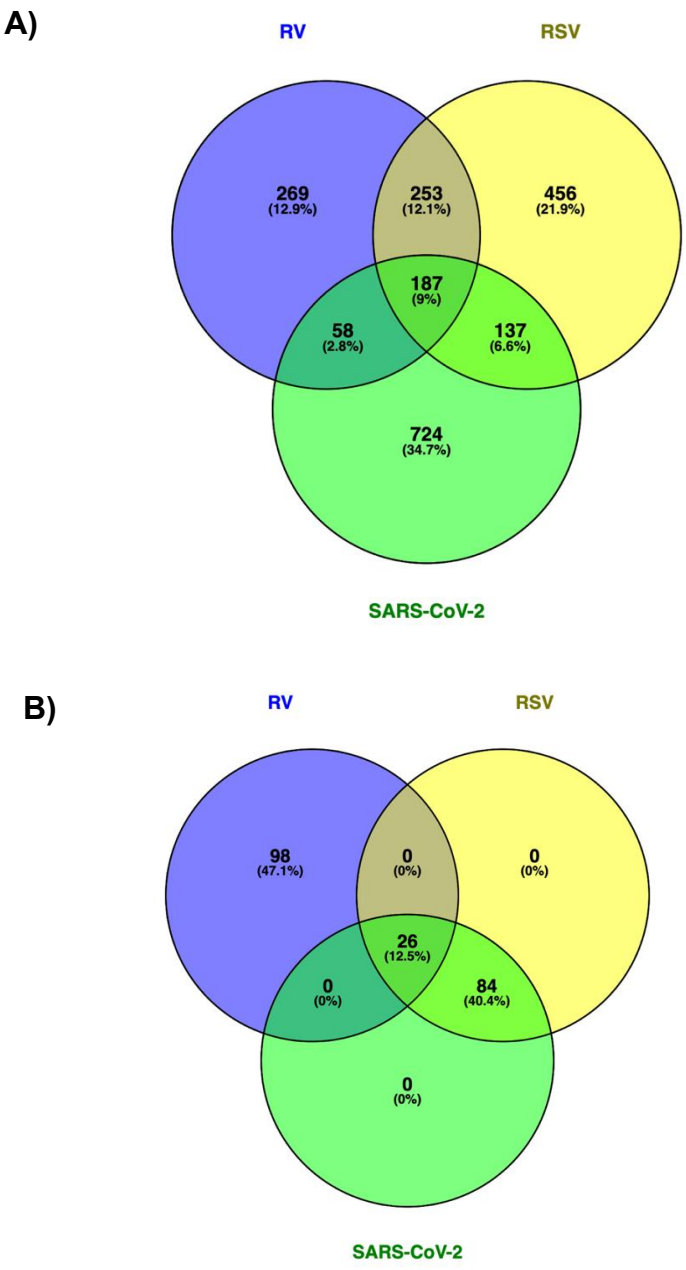


Figure 5-31. Venn diagram of overlapping pathways between RV, RSV and SARS-CoV-2. A) Pathways enriched for upregulated genes. **B)** Pathways enriched for downregulated genes. 187 upregulated pathways and 26 downregulated pathways were common between the three viruses.

Comparative analysis between Chapter 4 prolonged RV-A16 dataset and the published dataset by Bai *et al* was conducted. Bai *et al* utilised ALI to generate a differentiated mucociliary model of the bronchial airway epithelium. Whereas the Chapter 4 prolonged RV-A16 dataset employed submerged cultures of HBECs. 92 upregulated genes were common between the two dataset (Figure 5-32A). Moreover, 2306 and 365 upregulated genes were unique to Chapter 4 dataset and Bai *et al* dataset respectively. Genes encoding anti-viral restriction factors such as interferon-induced transmembrane protein 2 (IFITM2), OASL, OAS1 and OAS3 were common between the two dataset. IFITM2 is able to block early stages of viral replication[655]. Whereas OAS1 and OAS3 possess 2′-5′-oligoadenylate synthetase activity to degrade viral and cellular RNA to hinder viral replication[518]. Moreover, OASL mediates anti-viral innate immunity enhancing immunity to overcome viral evasion[656]. Genes implicated in a diverse range of function were found to be unique to the two dataset. These include genes implicated in cell cycle regulation (HORMAD1), calcium ion binding (CER1), and protein homodimerisation (GSTM1) which were unique to Chapter 4 dataset[518]. Whereas genes implicated in RNA degradation such as 5′-3′ exoribonuclease 1 (XRN1) were unique to Bai *et al* dataset[657].

160 downregulated genes were common between the two datasets (Figure 5-32B). Lysine demethylase 5B (KDM5B), which is involved in the regulation of cell differentiation, was one of the overlapping gene[658]. Another overlapping genes, deleted in azoospermia-associated protein 2 (DAZAP2), is implicated in DNA damage-induced p53 response. 1969 and 867 downregulated genes were found to be common between the two dataset[659]. Cyclase-associated protein 1 (CAP1), which mediates the dynamics of the actin cytoskeleton, was found to be unique to Chapter 4 dataset[660]. Whereas one of the downregulated genes unique to Bai *et al* dataset is ceramide kinase (CERK). CERK is involved in the catalysis of ceramide conversion to ceramide 1-phosphate[661].

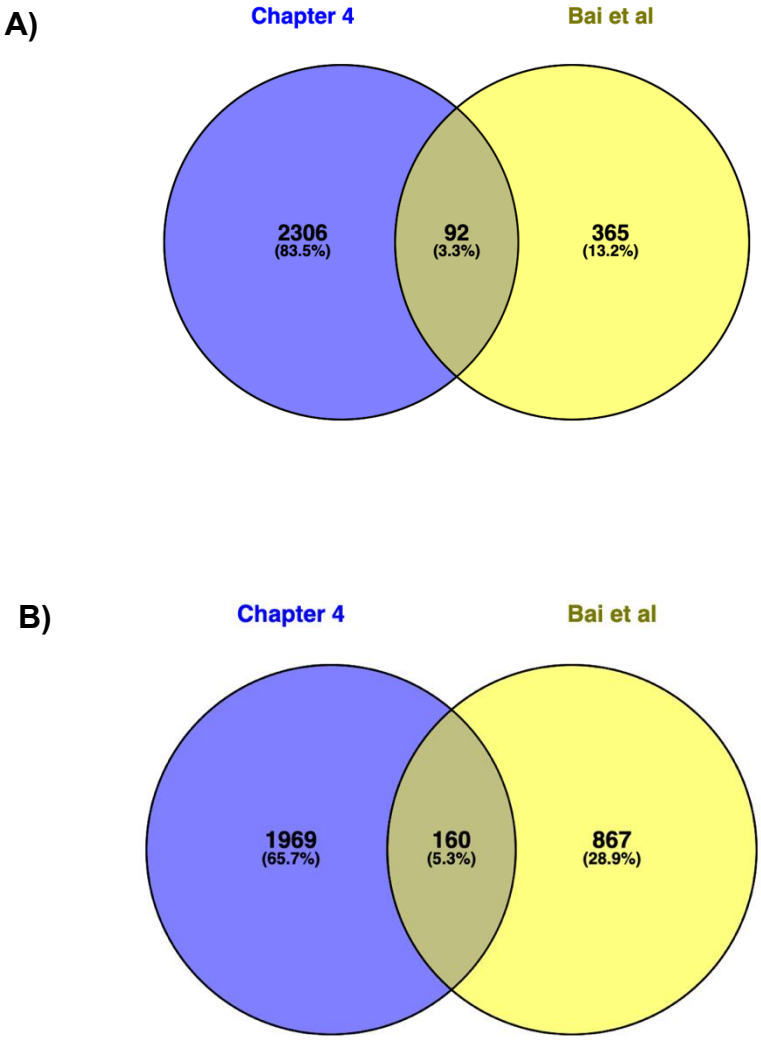


Figure 5-32. Venn diagram of overlapping genes Chapter 4 Prolonged RV-A16 dataset and published Bai *et al* dataset A) Upregulated. **B)** Downregulated. 92 upregulated genes and 160 downregulated genes were common between the two dataset.

Chapter 4 prolonged RSV-A2 dataset were compared with published dataset by Anderson *et al*. Similar to Bai *et al*, Anderson *et al* used ALI-HBEC model of the bronchial airway epithelium. 4998 were unique to Chapter 4 prolonged RSV-A2 dataset whereas 503 upregulated genes were unique to Anderson *et al* dataset (Figure 5-33A). 55 upregulated genes which are involved in wide range of functions were found to overlap between the two dataset. These include genes involved in antigen processing and presentation (HLA-DRB1)[518], neutrophil migration (intercellular adhesion molecule 2; ICAM2)[662], and purine nucleotide homeostasis (guanosine 5'-monophosphate reductase; GMPR)[663]. Upregulated genes unique to Chapter 4 prolonged RSV-A2 dataset include NEFM which is involved in cytoskeleton structure, ZIC5 which mediates DNA-binding transcription factor activity, and ABCB1 which is implicated in ATPase-coupled transmembrane transporter activity[518]. In contrast, upregulated genes unique to Anderson *et al* dataset include cytidine/uridine monophosphate kinase 2 (CMPK2) which facilitates phosphorylation of nucleotides[664], and bone marrow stromal antigen 2 (BST2) which is involved in hematopoietic stem cell activation[665].

72 downregulated gene were common between the two datsets (Figure 5-33B). Vacuolar protein sorting 41 (VPS41), which has an important role in the trafficking of lysosomal membrane proteins, is one of the overlapping downregulated genes[518]. AGO1 is another overlapping downregulated gene. Argonaute RISC component 1 (AGO1) is involved in microRNA-induced silencing complex[666]. In contrast, 6059 and 136 downregulated genes were unique for Chapter 4 prolonged RSV-A2 dataset and Anderson *et al* dataset. IPO7, which is involved in chronic inflammatory process of dental pulp, was one of the genes unique to Chapter 4 dataset[667]. Whereas secreted frizzled-related protein 4 (SFRP4), with roles in embryonic development, is unique to Anderson *et al* dataset[668]. It is important to note that utilising data from different labs which employ different experimental methodology is a limitation to the analysis. Nevertheless comparative analysis between published RV, RSV and SARS-CoV-2 dataset, and between Chapter 4 dataset gives insight into overlapping and unique genes and pathways furthering our understanding of host-virus interactions.

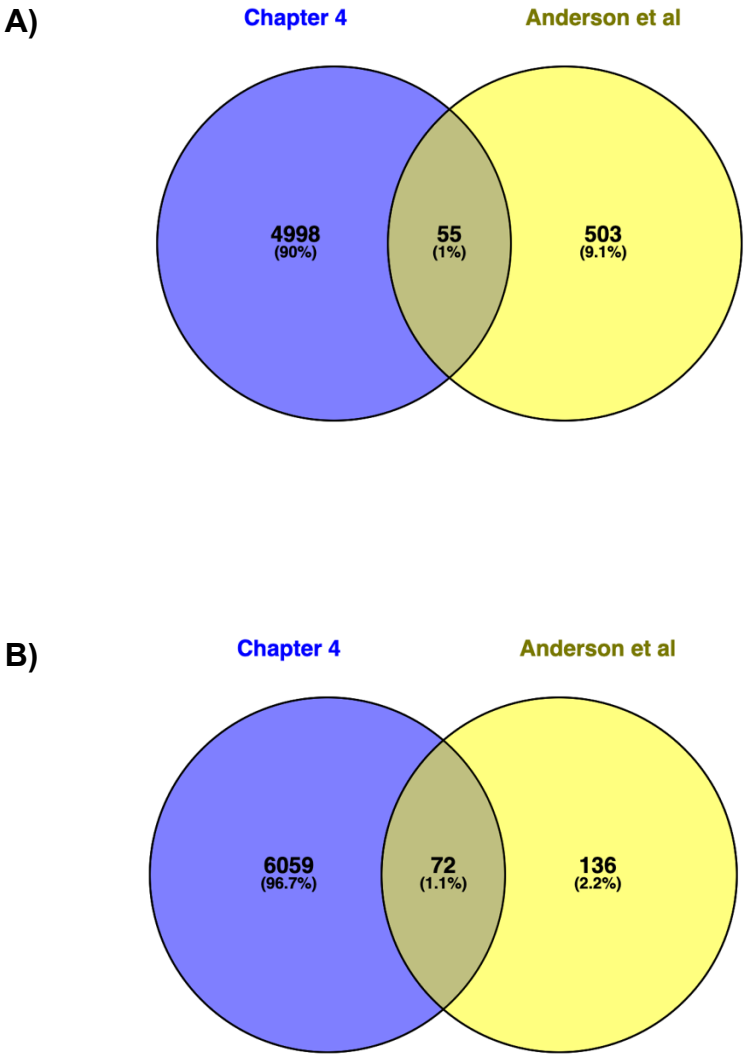


Figure 5-33. Venn diagram of overlapping genes Chapter 4 Prolonged RSV-A2 dataset and published Anderson *et al* dataset A) Upregulated. **B)** Downregulated. 55 upregulated genes and 72 downregulated genes were common between the two dataset.

5.5 Discussion

The main aims of the Chapter was to develop iPSC models of the airway epithelium and compare it to the ALI-HBEC model to investigate host-RSV interactions. To achieve these aims I utilised the Wong *et al* protocol to differentiate the iPSC cell line LOPCK and REBL-PAT into lung progenitors and mature airway epithelium. The differentiation of cells across the different stages was characterised by assaying for FOXA2 (Stage 1), SOX17 (Stage 1), NKX2.1 (Stage 2), FOXJ1 (Stage 3-4), KRT5 (Stage 3-4) and MUC5AC (Stage 3-4). DE and AFE were established in LOPCK cells whereas only DE was generated in REBL-PAT cells. This suggests that the LOPCK cell line may be better for reproducing Wong *et al*'s protocol. However both LOPCK and REBL-PAT were unsuccessful in differentiating into a mature airway epithelium whereas HBECs were successfully differentiated into mucociliary airway epithelium. This suggests issues with reproducibility of Wong *et al*'s protocol if different cell lines.

To investigate host-virus interactions, the utility of the model to study RSV infection was examined by assessing how permissive the iPSCs are, and by examining the changes in the expression of RSV receptor CX3CR1, interferon responses, and restriction factors OAS1 across key stages. RSV-A2 infection led to statistically significant elevation of CX3CR1 receptor and the OAS1 restriction factor. IFNβ response was detected in submerged and ALI-HBECs following RSV-A2 infection. Whereas IFNƛ was detected in submerged HBECs upon RSV-A2 infection. The detection of CX3CR1, IFNβ and OAS1 varied between non-infected and infected LOPCK and REBL-PAT cells in different stages of differentiation. However, this was found to be statistically insignificant suggesting RSV-A2 infection doesn't influence the expression of the receptor and interferon response in LOPCK and REBL-PAT cells. Preliminary experiments were conducted where LOPCK cells were investigated to see how permissive they were to RSV-A2 GFP by using flow cytometry to detect the mean percentage of GFP positive cells. Although the mean percentage of GFP positive cells increased as the stages of LOPCK differentiation increased, care must be taken when interpreting the results as the increase in GFP could be due to autofluorescence.

The expression of FOXA2 and SOX17 suggested the formation of DE in LOPCK and REBL-PAT cells. FOXA2 is a master regulator for the formation of DE[669]. In addition, development of organs of endodermal origin such as lungs, pancreas, and liver requires FOXA2[670]. Whereas SOX17 is a key transcription factor for DE specification and is used to confirm DE identity[671]. NKX2.1 expression in LOPCK suggested the generation of AFE. During embryonic lung development, NKX2.1 is involved in the differentiation of respiratory epithelial cells, lung morphogenesis and transcription of surfactant proteins SP-A, SP-B and SP-C, and secretary proteins from club cells[672]. However, NKX2.1 was not detected during Stage 2 in REBL-PAT cells. This suggests that the protocol by Wong *et al* could be more replicable in LOPCK cells than REBL-PAT cells.

The ciliated marker FOXJ1 was detected in Stage 1, 2 and 3 of LOPCK and REBL-PAT differentiation. This suggests that spontaneous differentiation may have occurred. Spontaneous differentiation can occur when cells located along the rim of the cell colony lack cell-to-cell contact at one open end, which could trigger uneven segregation during mitosis[673]. Furthermore, various somatic cells that are used to generate the iPSC clones have a variety of genetic and epigenetic profile[673]. This can influence the differentiation potential of the cells[674]. The expression of FOXJ1, KRT5 and MUC5AC in ALI-HBECs suggests that HBEC ALIs have differentiated into airway epithelium complete with basal, ciliated cells, and mucus-producing goblet cells. In contrast, the non-detection of FOXJ1, KRT5 and MUC5AC in stage 3 and 4 LOPCK and REBL-PAT cells indicates the undifferentiated state of these cells. This suggests that the ability to form specialised cells may be hindered by the immaturity of the cells. Surprisingly KRT5 was not detected in submerged HBECs. Keratin 5, a marker for basal cell airway progenitors, forms parallel coiled-coil heterodimers and assembles into a dynamic cytoskeleton to offer strength and flexibility to basal keratinocytes[675]. This could be due to low concentration of RNA used for RT-qPCR.

HBECs were successfully differentiated into pseudostratified epithelium with basal, ciliated and mucus producing goblet cells. The TEER readings increased as the days in ALI culture increased. Culturing cells for a longer time results in HBECs developing better tight junctions and stronger barrier integrity. Observation of mucus swirls and ciliated cells suggests that the HBECs had differentiated into a mucociliary layer. Nucleus, extracellular matrix and cytoplasm was visualised using H&E staining the mucus using Alcian-blue PAS staining. Haemotoxylin is a basic dye with an affinity for negatively charged structures such as DNA[676]. This gives a blue or basophilic appearance to the nucleus. Whereas eosin is an acidic dye which binds to positively charged structures such as cytoplasm to present a pink or eosinophilic appearance[676]. A catatonic dye, Alcian blue binds to negatively charged macromolecules such as mucopolysaccharides[677].

The attempt to differentiate and mature iPSCs into airway epithelium to investigate host-virus interaction was unsuccessful. One of the reasons for this could be due to the cell line. As one of the objectives was to replicate the protocol by Wong *et al* in different cell lines, LOPCK and REBL-PAT were used. The protocol by Wong *et al* was optimised for embryonic stem cell (ESC) line CA1[610]. Although the same genomic sequence is shared by somatic cells of an organism, subsequent generation of cells can obtain differential gene expression patterns after development from a single-cell zygote[678]. Through epigenetic modification, the cell specification process is reversed after somatic cells are reprogrammed to a pluripotent state[678]. This leads to re-establishment of the embryonic methylome and erasure of tissue-specific DNA methylation[678]. However, a study by Ghosh *et al* found that iPSCs can retain residual patterns of gene expression of their somatic cell type of origin. This epigenetic memory can lead to variation in differentiation capacity of different cell lines[679]. Whereas Carcamo-Orive *et al* reported that transcriptional variation between donors can influence the ability of cell lines to differentiate into specific cell types[680]. This can influence the capacity of the cells to respond to specific cell culture conditions[681]. Although a powerful tool with a wide range of applications, the differences between cell lines are limiting as differentiation protocols appear to work efficiently only in specific iPSC lines. Furthermore,

this means that replicating results and integrating data across different research groups remain challenging[682]. Another reason for the unsuccessful differentiation of the iPSCs into mature airway epithelium could be due to the quality of the reagents. Lot-to-lot variability could lead to reagents displaying variation in performance characteristics such as altered binding affinity, purity and specificity. This can hinder the quality of the reagents and potentially affect the differentiation capacity of the iPSC. Reagent quality could be tested by differentiating a small number of samples using different lots simultaneously to assess whether quality of the reagents is affecting the differentiation.

The iPSC and HBEC-derived models were developed and optimised to investigate host-virus interactions. One of the objectives was to assess whether LOPCK and REBL-PAT cells were permissive to RSV-A2 infection at different stages of differentiation and to compare them with submerged and HBEC ALIs. RSV-A2 was not detected by RT-qPCR in the virus-infected LOPCK and REBL-PAT cells. There was variation in changes of CX3CR1, IFNβ, and OAS1 between non-infected and infected LOPCK and REBL-PAT cells. However, comparison across groups using two-way ANOVA suggested no significant differences. This suggests that RSV-A2 may not influence the expression of CX3CR1, IFNβ, and OAS1 in LOPCK and REBL-PAT cells. Although RSV-A2 was not detected in submerged or HBEC ALIs, the detection of interferon response and OAS1 suggest that the cells may have been infected by the viruses, but the viral load is too low for detection. IFNβ was detected in submerged and HBEC ALIs following RSV-A2 infection. This is in line with literature as a study by Gortizka *et al* found that type I IFN production during RSV infection induces an anti-viral state whilst promoting pro-inflammatory response in the respiratory tract[683].

Following RSV-A2 infection, the expression of CX3CR1 and OAS1 was significantly elevated in submerged and HBEC ALIs. This is in line with literature as a study by Jeong *et al* found that CX3CR1 co-localises with RSV in a G protein-dependent manner in airway epithelium cells[684]. Upon infection, CX3CR1 binds to the conserved CX3C motif on the ectodomain RSV attachment G protein[685]. The interaction between RSV G protein and CX3CR1 is important in infection-induced responses of the airway epithelium including

production of cytokines such as IL8 and RANTES[686]. Upon viral infection, OAS1 is upregulated by IFNβ and IFNƛ [687]. OAS1 recognises and binds to viral dsRNA which stimulates the dimerization and activation of latent endoribonuclease L (RNase L)[324]. Following the binding of 2-5A and RNase L activation, the resulting viral and cellular RNA is degraded, and translation is inhibited[688]. The upregulation of OAS1 expression following RSV infection is submerged and HBEC ALIs suggests that OAS1 could potentially be used as a biomarker and/or therapeutic target for RSV infection.

Preliminary experiments were conducted in LOPCK cells where flow cytometry was employed to identify the percentage of RSV-infected cells by detecting GFP tagged to RSV-A2. The results showed that the percentage of GFP-positive cells increased as the stages of LOPCK iPSC differentiation progressed suggesting that the cells become more permissive as stages increase. The changes in the mean percentage of GFP cells between each stage of LOPCK differentiation and their infected counterpart was also found to be statistically significant. However care must be taken when interpreting results as although GFP was detected using flow cytometry, fluorescence was not observed using a microscope. This could potentially be attributed to the high sensitivity of flow cytometry in detecting fluorescence[674]. Whereas on the other hand, the increase in GFP detection could be due to autofluorescence[689]. It is important to consider that since green fluorescence was not seen under the microscope when cell lines were infected with RSV GFP, it is possible that fluorescence is not there and so the flow cytometry results may be false-positive. Due to time constraints, further investigation could not be carried out. The results should be considered preliminary and further studies are necessary for further validation before any meaningful conclusions can be made.

The strengths of the Chapter include using two different iPSC cell lines to reproduce the Wong *et al* protocol. This allows comparison of which iPSC cell is better for replicating the protocol. Reproducibility of iPSC protocols remains an issue. Differences between donor individuals can influence cellular traits of iPSCs such as pluripotency, differentiation, mRNA, DNA methylation and cell morphology[690]. This can lead to high variability in the differentiation potential

and the genetic stability between different iPSC cell lines[691]. Substantial experimental heterogeneity is an issue as attempts at experimental replication can be hindered by variation in derived differentiated cells[690]. Small variations in each step of the multi-step iPSC culture and differentiation protocol can accumulate and result in significantly different outcomes[692]. Rigorous quality control measures are imperative to standardise iPSC protocols to ensure that they are reproducible and meaningful[690]. Therefore it is important to replicate iPSC protocols in different iPSC cells. Another strength of the Chapter is the comparison of iPSC and ALI HBEC models. ALI HBEC models remain the gold standard for modelling the bronchial airway epithelium[10]. However, limitations of the model have led to the exploration of alternative models of the bronchial airway epithelium. Therefore comparing and contrasting the two models and using them to study RSV-A2 infection remains a strength of the Chapter.

Limitations of the Chapter include the need for purification strategies during the iPSC differentiation. The protocol by Wong *et al* was chosen for differentiation as it utilised the use of ALI technique to differentiate the iPSCs into mature bronchial airway epithelium. However, the protocol doesn't have methods to purify the cells at each stage of the differentiation. Therefore the iPSC cultures may potentially have mixed populations of non-lung cells. Through purification, cells expressing certain markers can be isolated from a mixed population of cells[693]. One strategy is the use of Fluorescence-Activated Cell Sorting (FACS) where antibodies targeting a specific cell surface protein is used to isolate cells[693]. Another promising strategy is Magnetic-Activated Cell Sorting (MACS) to collect antigen-positive target cells[694]. A study by Tsujisaka *et al* developed a purification technique which combined MACS and microRNA switch[694]. This technique separates and sorts target cells from heterogeneous cell populations through the use of flow cytometer[694]. Another limitation is that non-lung markers were not assayed due to time constraints. Assaying for non-lung markers can help assess whether the iPSCs are differentiating into different lineages such as pancreas or the thyroid. Therefore for future experiments a wide range of panel of markers should be used and cells be purified using the aforementioned purified techniques. Alternatively

RNA-seq could be used to measure the changes in gene expression levels across different stages of the differentiation.

Overall, LOPCK and REBL-PAT were differentiated into immature lung cells using Wong *et al* protocol. Although they were not able to form a mature pseudostratified epithelium, they nevertheless were valuable in investigating host-RSV-A2 interaction at different stages of differentiation. iPSC provides great promise for advancing the field of regenerative medicine. Furthermore, understanding the variability in iPSC protocol will enable strategies to be proposed to improve reproducibility of the protocols. iPSCs provide a great promise for advancing the field of regenerative medicine.

6.0 General Discussion

6.1 Aims and Objectives

The overall aim of this thesis was to gain greater understanding of host-virus interactions by using a genetic and transcriptomic- driven approach and applying this to cell biology techniques. One of the main hypotheses and aims were that genomics will allow us to understand host-virus interactions. Pathogenic viruses utilise host factors to disarm the host immune response to enable survival[695]. Genomic approaches to examine host-virus interactions have been increasingly employed to identify mechanisms underlying pathogenesis of viruses. Several studies investigating RV, RSV and SARS-CoV-2 infection have identified genetic signals associated with disease susceptibility and severity. However, it remains challenging to accurately pinpoint causal variants and the genes even if the locus has been identified[370]. Therefore, accurate determination of the causal genes from genetic signals is important[370]. I hypothesised that by translating genetic signals associated with RV, RSV and SARS-CoV-2 susceptibility and disease severity using a signal-to-gene pipeline, candidate causal genes can be identified and insight into pathways relevant to the viruses can be gained through downstream analysis. In addition, I aimed to identify upstream regulators associated with the three viruses to help uncover molecular events that may drive downstream changes. Genetic-driven approach has also been increasingly used to identify molecules as suitable biological targets for drug development[493]. Therefore, I aimed to identify potential druggable targets to enable the development of novel therapeutics by interrogating datasets.

Another aim of the thesis was to use transcriptomics to further understand host-virus interactions. Transcriptomics has been revolutionary in the field of respiratory medicine and has strengthened our understanding of host-virus interactions. RNA-seq has been utilised to profile the nasal airway epithelium, the airway and blood monocytes following influenza, hMPV and hRSV infection to reveal downregulation in cilium-related gene expression[499]. Moreover, the airway and blood monocyte transcriptome of individuals with severe RSV infection has been profiled by Chappin *et al* to identify differentially expressed

genes[503]. Although several studies have investigated the influence of acute respiratory viral infection on the host, little was known about prolonged infection. Therefore, I aimed to identify differentially expressed genes in response to prolonged RSV-A2 and RV-A16 infection through the use of RNA-Seq. Downstream analysis on the differentially expressed genes was conducted to identify and compare pathways and upstream regulators between the two viruses. I also aimed to investigate whether the viral gene signatures obtained from transcriptome profiling could provide support for genes identified from the signal-to-gene association analysis.

One of the aims of the thesis was to differentiate the iPSC cell lines LOPCK and REBL-PAT into lung progenitors and the mature airway epithelium using the Wong *et al* protocol and to compare and contrast them with ALI-HBEC model of the airway epithelium. Although ALI-HBEC model is the gold standard method for modelling the airway epithelium due to its ability to recapitulate the pseudostratified mucociliary epithelium, limitations such as finite supply of cells and the invasive nature used to extract cells has led to the development of alternative methods[10,594]. The unlimited capacity of expansion of iPSCs and the amenable nature for genetic engineering has made iPSC an attractive candidate[355]. Several iPSC protocols had been developed that enabled differentiation of IPSCs to airway epithelium[696]. The protocol by Wong *et al* employed the ALI technique to establish the airway epithelium[610]. Therefore the protocol was used to compare iPSC differentiation into airway epithelium with the ALI-HBEC model. I aimed to evaluate the differentiation of the iPSCs across different stages including Stage 1 DE, Stage 2 AFE, Stage 3 bronchial progenitors and Stage 4 expansion of bronchial progenitors. I also aimed to utilise the iPSC and ALI-HBEC models to study RSV infection by examining how permissive the cells were to iPSCs and then assessing the changes in the expression of RSV receptor CX3CR1, interferon response and the restriction factor OAS1 across different stages of differentiation.

6.2 Main Findings

Signal-to-gene pipeline analysis through interrogation of multiple datasets identified 6, 108 and 119 candidate causal genes associated with RV, RSV and SARS-CoV-2 respectively. Only 24 candidate causal genes were common for in more than one host-virus interactions with IL6, IL6-AS1 and STEAP1B showing overlap between all three viruses. RV has shown to be a potent stimulator for IL6 elaboration, and this RV-induced stimulation is mediated by NF-kB-mediated signalling pathway[583]. Whereas Pyle *et al* found that IL6 signalling during RSV infection stimulated IL27 dependent maturation of regulatory T cells to suppress inflammation[191]. A study by Santa Cruz *et al* identified IL6 as a biomarker for the development of fatal SARS-CoV-2 pneumonia following infection[697]. A significant negative correlation between peripheral oxygen saturation (SpO_2) and partial pressure of oxygen in arterial blood (PaO_2) was identified suggesting that IL6 is correlated with respiratory failure[697]. In comparison to IL6, literature focusing on IL6-AS1 and STEAP1B is low therefore further studies are needed to investigate the two genes.

Downstream analysis of candidate causal genes revealed that pathways implicating the modulation of the immune system may be particularly important between host RV/RSV/SARS-CoV-2 pathways. Th1 was the top pathway for RV. This is in line with literature as Th1 response is amplified during RV infection in individuals with allergic asthma[459]. Whereas macrophage classical activation signalling pathway was the top pathway for RSV and the IL10 signalling for SARS-CoV-2. RSV-challenged lung macrophage cells resulted in the upregulation of genes implicated in macrophage activation[461]. IL10 signalling was the top pathway for SARS-CoV-2. Early production of L10 has been found to be significantly associated with COVID-19 disease severity and may be useful as a prognostic biomarker to facilitate treatment strategies[463]. Pathways involving Th1 and Th2 activation were found to overlap between the three viruses. Diseases implicated in cellular development, connective tissue disorders, and infectious diseases were the top diseases for RV, RSV and SARS-CoV-2 respectively. Diseases affecting multiple different organ systems such as endocrine, nervous system and the gastrointestinal system were identified for SARS-CoV-2 in contrast to RV and RSV. This is in line with

literature as SARS-CoV-2 infection has shown to be associated with gastrointestinal and neurological symptoms[473,477]. Whereas COVID-19 has been suggested to induce diabetic ketoacidosis in individuals with diabetes[471].

Candidate causal genes were assessed to see whether they could be potential druggable targets for drug development or drug repurposing. It is essential to develop therapeutics for RV, RSV and SARS-CoV-2 infection. Using databases, drugs targeting 1, 16 and 11 candidate causal genes were currently in phase II, III and IV of clinical trials for RV, RSV and SARS-CoV-2 respectively. This included drugs targeting IL6 such as siltuximab, olokizumab and clazakizumab, and IC14 targeting CD14. Clazakizumab, which is used to treat psoriatic arthritis, was repurposed and found to significantly improve clinical outcomes, 28-day ventilator-free survival, 28- and 60- day overall survival in individuals hospitalised with severe or critical COVID-19 disease with hyperinflammation[495]. The role of genetics in the development of drugs is supported by literature. A study by Cook *et al* revealed that 73% drug discovery projects with genetic linkage were successful or active in Phase II trials compared with 43% of projects without the support[698]. Furthermore, genomic-driven approach to identify repurposable drug targets is an attractive proposition as it allows the use of de-risked compounds which could potentially decrease the developmental timelines and the overall developmental costs[699].

Transcriptomics was used to further understand host-virus interactions by studying how HBECs respond to prolonged RSV-A2 and RV-A16 infection. Overlapping upregulated genes such as XIST, CXCR4 and GSTM1 and overlapping downregulated genes including follistatin were identified. Pathways implicating the cell cycle may be particularly important for prolonged RSV-A2 infection. A study by Wu *et al* showed that cell cycle arrest at the G_0/G_1 phase can be induced by RSV infection after 24 hours[559]. Whereas the stimulation of DNA damage and stress response can result in the cell cycle checkpoints activity following 24 hour of RSV infection[560]. This suggests that RSV may utilise cell cycle checkpoints during acute and prolonged infection. In contrast, the top pathway with predicted significant activation was keratinisation. Pathways involving stress, apoptosis and cytoskeletal reorganisation may be important during prolonged RV-A16 infection. For

prolonged RSV-A2 infection, pathways implicating the inflammatory response were found to be inhibited with neutrophil degranulation being the top pathway with predicted inhibition. A study by Kirsebom *et al* observed that neutrophil activation peaks at 18-24 hours post-RSV infection which suggests that degranulation may transpire during acute stages of infection[565]. Neutrophil degranulation may become downregulated as RSV infection progresses. In addition, pathways involving formation of elastic fibres and idiopathic pulmonary fibrosis were found to be common between prolonged RSV-A2 and RV-A16 infection.

The viral gene signature was interrogated to see whether they provided support for genes supported by the signal-to-gene association analysis. Out of 119 candidate causal genes identified from the analysis, 28 were found to be differentially expressed in the acute dataset. These included genes involved in the immune system such as IFNAR1, IFNAR2, OAS1 and NFKBIZ. Significant increase in IFNAR1 concentration was reported by Yaugel-Novoa *et al* in individuals with severe COVID-19[586]. Whereas an inverse correlation was found between IRNAR1 concentration and severe COVID-19[586]. Analysis of prolonged RSV-A2 infection revealed that 43 genes out of 108 identified from signal-to-gene analysis were shown to be differentially expressed *in vitro*. Genes involved in a wide variety of roles such as immune response (CCR1), developmental processes (BMP4)[427], and ketogenesis (HMGCL)[547] were identified. For prolonged RV-A16 infection, 2 out of 6 identified genes were found to be differentially expressed. IL6 and STEAP1B, which was common for both RV and RSV, was found to be significantly downregulated during prolonged RV and RSV infection. This is supported by literature where bioactivity of IL6 has been shown to peak at 48-72 hours following RV infection and reducing thereafter[583].

To compare and contrast iPSC and ALI HBEC models of the airway epithelium, protocol by Wong *et al* was utilised to differentiate the LOPCK and REBL-PAT cell line into bronchial progenitors and the airway epithelium. The formation of DE and AFE in LOPCK cells but only the generation of DE in REBL-PAT cells suggests that the Wong *et al* protocol may be more reproducible in LOPCK cells. The non-detection of FOXJ1, KRT5 and MUC5AC in Stage 3 and 4

LOPCK and REBL-PAT cells could be due to the immaturity of the cells hindering the ability to form specialised cells. However, mature airway epithelium could not be established in LOPCK and REBL-PAT iPSCs. Whereas HBECs cultured in ALI were successfully differentiated into a mucociliary airway epithelium. The unsuccessful attempts at differentiating the iPSCs into the airway epithelium could be due to the cell line used since the Wong *et al* protocol was optimised for the CA1 embryonic stem cell line[610]. A study by Ortmann *et al* showed that cell lines derived from distinct genetic backgrounds can lead to divergence in differentiation capacity[681]. Whereas Carcamo-Drive *et al* found that iPSCs retain a donor specific gene expression pattern which can lead to variability across cell lines and influence differentiation potency [680]. Therefore protocols for iPSC differentiation appear to work efficiently in specific cell lines. These findings are important as rigorous and standardised iPSC protocols are needed. Strategies could be developed to improve reproducibility of iPSC differentiation protocols and data across different research groups could be integrated[690].

iPSCs and HBECs were utilised to study host-RSV interaction by assessing whether LOPCK, REBL-PAT, submerged and ALI HBECs were permissive to RSV-A2. Significant increase in RSV receptor CX3CR1 and the restriction factor OAS1 was found in RSV-infected submerged and ALI-HBECs compared to non-infected. Upon RSV infection, CX3CR1 has been shown to co-localise with RSV in airway epithelial cells[684]. Whereas OAS1 has shown to be upregulated following RSV infection in airway epithelial cells[700]. IFN-ƛ response was seen in RSV-infected submerged and ALI-HBECs whereas IFN-β response was present in RSV-infected submerged cells. During RSV infection, IFN response are amplified pro-inflammatory response to induce an anti-viral state[683]. Variability in CX3CR1, IFN-β and OAS1 was seen across different stages of differentiation in non-infected and infected LOPCK and REBL-PAT cells. The difference was non-significant suggesting RSV-A2 infection may not influence the expression of CX3CR1, IFN-β and OAS1 in LOPCK and REBL-PAT cells. However, preliminary experiments suggested that LOPCK cells may become increasingly permissive to RSV-A2 tagged with GFP since the mean percentage of GFP-positive cells increased in infected

LOPCK cells as the stages of differentiation progressed. However, care must be taken when interpreting the results as this increase in fluorescence could be due to autofluorescence. Therefore further investigation is needed to validate the results.

6.3 Study Limitations

One limitation of the study is that mature airway epithelium derived from iPSCs could not be established as the iPSC cells failed to proliferate. One of the reasons for this could be due to the cell line used. The Wong *et al* protocol was optimised for the CA1 cell line[610]. Variability among iPSCs can arise from epigenetic and genetic factors which can influence the behaviour of the cells[681]. This remains a prevailing issue as it can hinder the differentiation capacity of the cells and the experimental reproducibility[681]. Another limitation of the study Wong *et al* protocol did not enable the purification of cells at different stages of the iPSC differentiation. Therefore the cell culture could be heterogenous with a lot of non-lung cells present. Non-lung markers were not assayed due to time constraints. This remains a limitation as this could help examine whether the cells are differentiating into different lineages.

Another limitation was the preliminary experiments used to study whether LOPCK cells were permissive to RSV-A2 infection at different stages of the differentiation. The RSV-A2 was tagged with GFP, and flow cytometry was used to measure the number of GFP positive cells. Although the mean percentage of GFP increased in infected cells as they progressed through differentiation, fluorescence of the infected cells were not observed under a microscope. This could be due to high sensitivity of flow cytometry in detecting fluorescent signals[639]. On the other hand the increase in fluorescence could be due to autofluorescence. However, time did not permit further investigation. Therefore further investigation is necessary for validation before any meaningful conclusions can be made.

Another study limitation is that many differentially expressed genes identified during prolonged infection appear to be related to cell homeostasis and cell death. It is important to consider that this could be potentially due to cells senescing or undergoing apoptosis after 6 days in culture. Another limitation is that acute and prolonged response between the RV, RSV and SARS-CoV-2. Differentially expressed genes in response to acute RV and RSV infection in submerged HBECs were not available in published literature. Therefore similarities and differences between acute and prolonged RV and RSV infection could not be made. Since prolonged infection of SARS-CoV-2 was

not conducted due to time constraints, changes in the transcriptomic profile between acute and prolonged SARS-CoV-2 could not be made.

6.4 Future Work

Future work would prioritise optimisation of Wong *et al* iPSC cell culture protocol such as optimising media formulations and incorporating methods to purify cells with specific markers at different stages of the differentiation. Fluorescence-Activated Cell Sorting (FACS) could be implemented to identify and isolate distinct cell types at the single-cell level[355]. FACS employs the use of cell surface antigens expressed by cells of interest to isolate cells. Another method to purify cells involves the targeting of endogenous miRNA activity[701]. Synthetic mRNAs called microRNA which encode fluorescent protein are tagged with sequences which are then targeted by miRNAs expressed on cells of interest[701]. Whereas Magnetic-Activated Cell Sorting (MACS), which uses superparamagnetic nanoparticles and columns to capture cells expressing certain markers, can be combined with microRNA switch method to purify large amounts of iPSCs rapidly and efficiently[694]. Furthermore, single-cell RNA sequencing could be used to distinguish heterogenous cell populations[690]. In addition, a wider panel of markers such as pancreatic, thyroid and intestine markers would also be used to assess whether the cells are differentiating into different lineages. Utilising RNA-seq would also be useful to examine the changes in the expression levels of the different markers as the cells progress through differentiation.

By optimising the Wong *et al* protocol to enable differentiation of iPSCs into a mature airway epithelium, candidate causal genes and pathways that have been identified through signal-to-gene analysis can be interrogated in both iPSC and ALI HBEC models. Clustered Regularly Interspaced Short Palindromic repeats (CRISPR)/Cas9 is a genome editing technology CRISPR-Cas9 that enables efficient and precise modification of DNA sequences[702]. The system uses specifically designed guide RNA to target a desired DNA sequence and Cas9 enzyme to recognise and cut the DNA at the specific location[702]. CRISPR-Cas9 can be employed to edit candidate casual genes identified from the signal-to-pathway analysis in iPSC and HBEC-derived ALI models of the bronchial airway epithelium. Further mechanistic and functional studies can be conducted to investigate the downstream effects of the candidate causal genes on RV, RSV and SARS-CoV-2 infection. In addition,

mechanistic studies on the identified pathways and upstream regulators can be conducted. This can give further insight into the molecular basis of viral pathogenesis.

To compare and contrast between acute and prolonged RV, RSV and SARS-CoV-2, future studies would prioritise the transcriptomic profiling of acute RV and RSV infection, and prolonged SARS-CoV-2 infection in submerged HBECs. Downstream analysis using IPA could be conducted to identify pathways and upstream regulators that are unique or common between acute and prolonged infection in the three viruses. This can give insight on the viral pathogenesis and how it can differ during acute and prolonged infection therefore providing further knowledge on the mechanistic background. In addition, biomarkers or druggable targets can be identified to guide treatment strategies to limit RV, RSV and SARS-CoV-2 infection. Pathways and upstream regulators identified from acute and prolonged response could also be modelled using iPSC and HBEC-derived ALI models to further investigate the molecular background.

6.5 Final Conclusions

Signal-to-gene analysis has provided significant understanding of genes, pathways, diseases and upstream regulators which may be particularly relevant for RV, RSV and SARS-CoV-2. Candidate causal genes were identified with IL6, IL6-AS1 and STEAP1B being common for all three viruses. Downstream analysis revealed that immunomodulatory pathways could be particularly important for all three viruses. Moreover, SARS-CoV-2 might be implicated in diseases affecting multiple organ systems compared to RV and RSV. Overlapping and unique pathways, diseases and upstream regulators were identified by studying the similarities and differences between the three viruses. Interrogation of downstream effects is particularly important as it builds on our understanding of viral pathogenesis. A significant opportunity has been provided by these findings for the development of new therapeutic strategies. Candidate causal genes may serve as biomarkers or druggable targets to limit viral infection. Studying the similarities and differences between the three viruses using this genetic-based approach can lead to effective repurposing of drugs that exhibit broad anti-viral activity.

Host-virus interactions were further investigated using transcriptomics to understand the response of HBECs upon prolonged RSV-A2 and RV-A16 infection. Cell cycle pathways might be important for prolonged RSV-A2 infection whereas keratinisation was the top pathway predicted to be significantly activated. Overlapping upstream regulators such as MEK1 inhibitor PD98059 were identified. These analyses can help generate hypotheses about pathways, disease mechanisms and regulatory networks which can guide future research. Viral gene signature offered some support to genes translated from signal-to-gene association analysis. Genes such as CCR1, BMP4 and HMGCL were differentially expressed following prolonged RSV-A2 *in vitro*. Interrogation of acute SARS-CoV-2 dataset provided further understanding of host-SARS-CoV-2 interactions.

Since the advent of iPSC technology, advances in the field including differentiation of lung, cardiac, neural and hepatic cell types have been documented[696]. Although iPSC-derived airway epithelium could not be established, iPSCs were still useful in studying host-virus interactions at

different stages of the differentiation. Furthermore, reproducing iPSC protocols in a variety of different cell lines is important for informing the wider research community on whether the methodology is replicable or not. This can lead to strategies being implemented to improve the reproducibility of the protocol. Volpato *et al* suggested that the development of standardised protocols with the documentation of yield, fate and purity of the derived cell types could elevate the reliability of phenotype comparisons and therefore improve reproducibility across different research groups[690]. Nevertheless, the potential of iPSCs in studying host-virus interactions and disease modelling is vast. Currently the next step towards creating complex 'mini lung' models have focused on combining the airway epithelium cells with non-epithelial cell types such as endothelial, stromal and immune cells[703]. Developing reproducible protocols that capture this complex architecture of the airway epithelium has potential to further our understanding of host-virus interactions.

Overall, the use of genetic and transcriptomic-based approaches has identified genes, pathways, and upstream regulators that are relevant to RV, RSV and SARS-CoV-2. Opportunities for the development of therapeutics have been provided by these findings. Although iPSC-derived models of the airway epithelium could not be established, the use of cell models has been useful for studying host-RSV interactions. This thesis has led to understanding of pathways and diseases that underlie host-RV/RSV/SARS-CoV-2 interactions.

7.0 Bibliography

1 Davis JD, Wypych TP. Cellular and functional heterogeneity of the airway epithelium. Mucosal Immunol. 2021; **14**. DOI:10.1038/s41385-020-00370-7.

2 Ball M, Hossain M, Padalia D. Anatomy Airway. In*:* StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing 2022.

3 Patwa A, Shah A. Anatomy and physiology of respiratory system relevant to anaesthesia. Indian J. Anaesth. 2015; **59**. DOI:10.4103/0019-5049.165849.

4 Otter CJ, Fausto A, Tan LH, Khosla AS, Cohen NA, Weiss SR. Infection of primary nasal epithelial cells differentiates among lethal and seasonal human coronaviruses. *Proc Natl Acad Sci U S A* 2023; **120**. DOI:10.1073/pnas.2218083120.

5 Farzal Z, Basu S, Burke A, *et al.* Comparative study of simulated nebulized and spray particle deposition in chronic rhinosinusitis patients. *Int Forum Allergy Rhinol* 2019; **9**. DOI:10.1002/alr.22324.

6 Brand-Saberi BEM, Schäfer T. Trachea: Anatomy and physiology. Thorac. Surg. Clin. 2014; **24**. DOI:10.1016/j.thorsurg.2013.09.004.

7 Lungova V, Thibeault SL. Mechanisms of larynx and vocal fold development and pathogenesis. Cellular and Molecular Life Sciences. 2020; **77**. DOI:10.1007/s00018-020-03506-x.

8 Amador C, Weber C, Varacallo M. Anatomy, Thorax, Bronchial. In*:* StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022.

9 Shaykhiev R. Multitasking basal cells: Combining stem cell and innate immune duties. In: European Respiratory Journal. 2015. DOI:10.1183/13993003.00521-2015.

10 Michi AN, Proud D. A toolbox for studying respiratory viral infections using air-liquid interface cultures of human airway epithelial cells. *Am J*

Physiol Lung Cell Mol Physiol 2021; **321**. DOI:10.1152/AJPLUNG.00141.2021.

11 Knight DA, Holgate ST. The airway epithelium: Structural and functional properties in health and disease. Respirology. 2003; **8**. DOI:10.1046/j.1440-1843.2003.00493.x.

12 Raby KL, Michaeloudes C, Tonkin J, Chung KF, Bhavsar PK. Mechanisms of airway epithelial injury and abnormal repair in asthma and COPD. Front. Immunol. 2023; **14**. DOI:10.3389/fimmu.2023.1201658.

13 Whitsett JA. Airway epithelial differentiation and mucociliary clearance. In: Annals of the American Thoracic Society. 2018. DOI:10.1513/AnnalsATS.201802-128AW.

14 Ma J, Rubin BK, Voynow JA. Mucins, Mucus, and Goblet Cells. Chest. 2018; **154**. DOI:10.1016/j.chest.2017.11.008.

15 Ganesan S, Comstock AT, Sajjan US. Barrier function of airway tract epithelium. *Tissue Barriers* 2013; **1**. DOI:10.4161/tisb.24997.

16 Zuo WL, Shenoy SA, Li S, *et al.* Ontogeny and biology of human small airway epithelial club cells. *Am J Respir Crit Care Med* 2018; **198**. DOI:10.1164/rccm.201710-2107OC.

17 Tam A, Wadsworth S, Dorscheid D, Man S f. P, Sin DD. The airway epithelium: More than just a structural barrier. Ther. Adv. Respir. Dis. 2011; **5**. DOI:10.1177/1753465810396539.

18 Hewitt RJ, Lloyd CM. Regulation of immune responses by the airway epithelial cell landscape. Nat. Rev. Immunol. 2021; **21**. DOI:10.1038/s41577-020-00477-9.

19 Ye Q, Bankova LG. Brush cells fine-tune neurogenic inflammation in the airways. *Journal of Clinical Investigation* 2022; **132**. DOI:10.1172/JCI161439.

20 Reid L, Meyrick B, Antony VB, Chang L-Y, Crapo JD, Reynolds HY. The Mysterious Pulmonary Brush Cell. *Am J Respir Crit Care Med* 2005; **172**. DOI:10.1164/rccm.200502-203ws.

21 Perniss A, Boonen B, Tonack S, *et al.* A succinate/SUCNR1-brush cell defense program in the tracheal epithelium. *Sci Adv* 2023; **9**. DOI:10.1126/sciadv.adg8842.

22 Gustafsson BI, Kidd M, Chan A, Malfertheiner M V., Modlin IM. Bronchopulmonary neuroendocrine tumors. Cancer. 2008; **113**. DOI:10.1002/cncr.23542.

23 Cutz E, Yeger H, Pan J. Pulmonary neuroendocrine cell system in pediatric lung disease - Recent advances. Pediatric and Developmental Pathology. 2007; **10**. DOI:10.2350/07-04-0267.1.

24 Adriaensen D, Scheuermann DW. Neuroendocrine cells and nerves of the lung. *Anat Rec* 1993; **236**. DOI:10.1002/ar.1092360111.

25 Montoro DT, Haber AL, Biton M, *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 2018; **560**. DOI:10.1038/s41586-018-0393-7.

26 Goldfarbmuren KC, Jackson ND, Sajuthi SP, *et al.* Dissecting the cellular specificity of smoking effects and reconstructing lineages in the human airway epithelium. *Nat Commun* 2020; **11**. DOI:10.1038/s41467-020-16239-z.

27 Sato Y, Kim D, Turner MJ, *et al.* Ionocyte-Specific Regulation of Cystic Fibrosis Transmembrane Conductance Regulator. *Am J Respir Cell Mol Biol* 2023; **69**. DOI:10.1165/rcmb.2022-0241OC.

28 Luan X, Romero NH, Campanucci VA, *et al.* Pulmonary Ionocytes Regulate Airway Surface Liquid pH in Primary Human Bronchial Epithelial Cells. *Am J Respir Crit Care Med* 2024. DOI:10.1164/rccm.202309-1565OC.

29 Giraud-Gatineau A, Colson P, Jimeno MT, *et al.* Comparison of mortality associated with respiratory viral infections between December

2019 and March 2020 with that of the previous year in Southeastern France. *International Journal of Infectious Diseases* 2020; **96**. DOI:10.1016/j.ijid.2020.05.001.

30 Shi T, Arnott A, Semogas I, *et al.* The etiological role of common respiratory viruses in acute respiratory infections in older adults: A systematic review and meta-analysis. *Journal of Infectious Diseases* 2020; **222**. DOI:10.1093/INFDIS/JIY662.

31 Troeger C, Forouzanfar M, Rao PC, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect Dis* 2017; **17**. DOI:10.1016/S1473-3099(17)30396-1.

32 Lynch JP, Kajon AE. Adenovirus: Epidemiology, Global Spread of Novel Serotypes, and Advances in Treatment and Prevention. *Semin Respir Crit Care Med* 2016; **37**. DOI:10.1055/s-0036-1584923.

33 Weston S, Frieman MB. Respiratory viruses. In: Encyclopedia of Microbiology. 2019. DOI:10.1016/B978-0-12-801238-3.66161-5.

34 Boncristiani H, Criado M, Arruda E. Respiratory Viruses. Encyclopedia of Microbiology. 2009. https://pmc.ncbi.nlm.nih.gov/articles/PMC7149556/.

35 Dodge MJ, MacNeil KM, Tessier TM, Weinberg JB, Mymryk JS. Emerging antiviral therapeutics for human adenovirus infection: Recent developments and novel strategies. Antiviral Res. 2021; **188**. DOI:10.1016/j.antiviral.2021.105034.

36 Yan Z, Deng X, Qiu J. Human Bocavirus 1 Infection of Well-Differentiated Human Airway Epithelium. *Curr Protoc Microbiol* 2020; **58**. DOI:10.1002/cpmc.107.

37 Lynch JP, Kajon AE. Adenovirus: Epidemiology, Global Spread of Novel Serotypes, and Advances in Treatment and Prevention. *Semin Respir Crit Care Med* 2016; **37**. DOI:10.1055/s-0036-1584923.

38 Nguyen S, Baker K, Padman BS, *et al.* Bacteriophage Transcytosis Provides a Mechanism To Cross Epithelial Cell Layers. *mBio* 2017. DOI:10.1128/mbio.01874-17.

39 Obenaus M, Schildgen O, Schürmann D, *et al.* High-dose glucocorticoid treatment of near-fatal bocavirus lung infection results in rapid recovery. ERJ Open Res. 2021; **7**. DOI:10.1183/23120541.00060-2021.

40 Javanian M, Barary M, Ghebrehewet S, Koppolu V, Vasigala VKR, Ebrahimpour S. A brief review of influenza virus infection. J. Med. Virol. 2021; **93**. DOI:10.1002/jmv.26990.

41 Chesnokov A, Ivashchenko AA, Matsuzaki Y, *et al.* Influenza C virus susceptibility to antivirals with different mechanisms of action. *Antimicrob Agents Chemother* 2024. https://pmc.ncbi.nlm.nih.gov/articles/PMC11064526/.

42 Paules C, Subbarao K. Influenza. The Lancet. 2017; **390**. DOI:10.1016/S0140-6736(17)30129-0.

43 Morens DM, Taubenberger JK. Making Universal Influenza Vaccines: Lessons from the 1918 Pandemic. *Journal of Infectious Diseases* 2019; **219**. DOI:10.1093/infdis/jiy728.

44 Taubenberger JK, Morens DM. The pathology of influenza virus infections. Annual Review of Pathology: Mechanisms of Disease. 2008; **3**. DOI:10.1146/annurev.pathmechdis.3.121806.154316.

45 Mertz D, Geraci J, Winkup J, Gessner BD, Ortiz JR, Loeb M. Pregnancy as a risk factor for severe outcomes from influenza virus infection: A systematic review and meta-analysis of observational studies. Vaccine. 2017; **35**. DOI:10.1016/j.vaccine.2016.12.012.

46 Kuo YC, Lai CC, Wang YH, Chen CH, Wang CY. Clinical efficacy and safety of baloxavir marboxil in the treatment of influenza: A systematic review and meta-analysis of randomized controlled trials. *Journal of Microbiology, Immunology and Infection* 2021; **54**. DOI:10.1016/j.jmii.2021.04.002.

47 Nadiger M, Sendi P, Martinez PA, Totapally BR. Epidemiology and Clinical Features of Human Metapneumovirus and Respiratory Syncytial Viral Infections in Children. *Pediatric Infectious Disease Journal* 2023; **42**. DOI:10.1097/INF.0000000000004055.

48 Henrickson KJ, Hoover S, Kehl KS, Hua W. National disease burden of respiratory viruses detected in children by polymerase chain reaction. In: Pediatric Infectious Disease Journal. 2004. DOI:10.1097/01.inf.0000108188.37237.48.

49 Jacobs SE, Lamson DM, Kirsten S, Walsh TJ. Human rhinoviruses. *Clin Microbiol Rev* 2013; **26**. DOI:10.1128/CMR.00077-12.

50 Liu L, Pan Y, Zhu Y, *et al.* Association between rhinovirus wheezing illness and the development of childhood asthma: A meta-analysis. *BMJ Open* 2017; **7**. DOI:10.1136/bmjopen-2016-013034.

51 Blaas D, Fuchs R. Mechanism of human rhinovirus infections. *Mol Cell Pediatr* 2016; **3**. DOI:10.1186/s40348-016-0049-3.

52 Waman VP, Kolekar PS, Kale MM, Kulkarni-Kale U. Population structure and evolution of rhinoviruses. *PLoS One* 2014; **9**. DOI:10.1371/journal.pone.0088981.

53 Palmenberg AC. Rhinovirus C, Asthma, and Cell Surface Expression of Virus Receptor CDHR3. *J Virol* 2017; **91**. DOI:10.1128/jvi.00072-17.

54 Lamson D, Renwick N, Kapoor V, *et al.* Masstag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004-2005. *Journal of Infectious Diseases* 2006; **194**. DOI:10.1086/508551.

55 Bochkov YA, Gern JE. Clinical and molecular features of human rhinovirus C. Microbes Infect. 2012; **14**. DOI:10.1016/j.micinf.2011.12.011.

56 Bochkov YA, Gern JE. Rhinoviruses and Their Receptors: Implications for Allergic Disease. Curr. Allergy Asthma Rep. 2016; **16**. DOI:10.1007/s11882-016-0608-7.

57 Warner SM, Wiehler S, Michi AN, Proud D. Rhinovirus replication and innate immunity in highly differentiated human airway epithelial cells. *Respir Res* 2019; **20**. DOI:10.1186/s12931-019-1120-0.

58 Triantafilou K, Vakakis E, Richer EAJ, Evans GL, Villiers JP, Triantafilou M. Human rhinovirus recognition in non-immune cells is mediated by toll-like receptors and MDA-5, which trigger a synergetic pro-inflammatory immune response. *Virulence* 2011; **2**. DOI:10.4161/viru.2.1.13807.

59 Vareille M, Kieninger E, Edwards MR, Regamey N. The airway epithelium: Soldier in the fight against respiratory viruses. *Clin Microbiol Rev* 2011; **24**. DOI:10.1128/CMR.00014-10.

60 Falsey AR, Branche AR. Rhinoviruses. In: International Encyclopedia of Public Health. Elsevier, 2017: 363–9.

61 Schilder AGM, Chonmaitree T, Cripps AW, *et al.* Otitis media. *Nat Rev Dis Primers* 2016; **2**: 16063.

62 Monasta L, Ronfani L, Marchetti F, *et al.* Burden of disease caused by otitis media: Systematic review and global estimates. PLoS One. 2012; **7**. DOI:10.1371/journal.pone.0036226.

63 Jaume F, Valls-Mateus M, Mullol J. Common Cold and Acute Rhinosinusitis: Up-to-Date Management in 2020. Curr. Allergy Asthma Rep. 2020; **20**. DOI:10.1007/s11882-020-00917-5.

64 Heikkinen T, Järvinen A. The common cold. *The Lancet* 2003; **361**: 51–9.

65 DeBoer DL, Kwon E. Acute Sinusitis. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024.

66 Mims JW. Asthma: Definitions and pathophysiology. *Int Forum Allergy Rhinol* 2015; **5**. DOI:10.1002/alr.21609.

67 Jackson DJ, Gangnon RE, Evans MD, *et al.* Wheezing rhinovirus illnesses in early life predict asthma development in high-risk children. *Am J Respir Crit Care Med* 2008; **178**. DOI:10.1164/rccm.200802-309OC.

68 Makris S, Johnston S. Recent advances in understanding rhinovirus immunity. F1000Res. 2018; **7**. DOI:10.12688/F1000RESEARCH.15337.1.

69 Berger A. Science commentary: Th1 and Th2 responses: What are they? Br. Med. J. 2000; **321**. DOI:10.1136/bmj.321.7258.424.

70 Braman SS. The global burden of asthma. In: Chest. 2006. DOI:10.1378/chest.130.1_suppl.4S.

71 McCracken JL, Veeranki SP, Ameredes BT, Calhoun WJ. Diagnosis and management of asthma in adults a review. JAMA - Journal of the American Medical Association. 2017; **318**. DOI:10.1001/jama.2017.8372.

72 Glanville N, Johnston SL. Challenges in developing a cross-serotype rhinovirus vaccine. Curr. Opin. Virol. 2015; **11**. DOI:10.1016/j.coviro.2015.03.004.

73 Higgins PG, Tyrrell DAJ, Barclay WS, al-Nakib W. The time course of the humoral immune response to rhinovirus infection. *Epidemiol Infect* 1989; **103**. DOI:10.1017/S095026880003106X.

74 Borchers AT, Chang C, Gershwin ME, Gershwin LJ. Respiratory syncytial virus - A comprehensive review. Clin. Rev. Allergy Immunol. 2013; **45**. DOI:10.1007/s12016-013-8368-9.

75 Barr R, Green CA, Sande CJ, Drysdale SB. Respiratory syncytial virus: diagnosis, prevention and management. *Ther Adv Infect Dis* 2019; **6**. DOI:10.1177/2049936119865798.

76 Chertow DS. SARS, MERS, and Other Viral Lung Infections. *Crit Care Med* 2017; **45**. DOI:10.1097/ccm.0000000000002420.

77 Simões EAF. The Outpatient Burden of Respiratory Syncytial Virus Infections in Older Children. J. Infect. Dis. 2017; **215**. DOI:10.1093/infdis/jiw483.

78 Falsey AR, Walsh EE. Respiratory syncytial virus infection in elderly adults. Drugs Aging. 2005; **22**. DOI:10.2165/00002512-200522070-00004.

79 Sigurs N, Aljassim F, Kjellman B, *et al.* Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life. *Thorax* 2010; **65**. DOI:10.1136/thx.2009.121582.

80 Smith DK, Seales S, Budzik C. Respiratory syncytial virus bronchiolitis in children. *Am Fam Physician* 2017; **95**.

81 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; **17**. DOI:10.1038/nm.2444.

82 Griffiths C, Drews SJ, Marchant DJ. Respiratory syncytial virus: Infection, detection, and new options for prevention and treatment. *Clin Microbiol Rev* 2017; **30**. DOI:10.1128/CMR.00010-16.

83 Justice NA, Le JK. Bronchiolitis. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024.

84 Bozzola E, Ciarlitto C, Guolo S, *et al.* Respiratory Syncytial Virus Bronchiolitis in Infancy: The Acute Hospitalization Cost. *Front Pediatr* 2021; **8**. DOI:10.3389/fped.2020.594898.

85 Lin HC, Liu YC, Hsing TY, *et al.* RSV pneumonia with or without bacterial co-infection among healthy children. *Journal of the Formosan Medical Association* 2022; **121**. DOI:10.1016/j.jfma.2021.08.012.

86 Shi T, McAllister DA, O'Brien KL, *et al.* Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *The Lancet* 2017; **390**. DOI:10.1016/S0140-6736(17)30938-8.

87 Kim T, Huh JW, Hong SB, *et al.* Epidemiology and Characteristics of Respiratory Syncytial Virus Pneumonia in Critically Ill Adults. *Open Forum Infect Dis* 2023; **10**. DOI:10.1093/ofid/ofad131.

88 Rosas-Salazar C, Chirkova T, Gebretsadik T, *et al.* Respiratory syncytial virus infection during infancy and asthma during childhood in the USA (INSPIRE): a population-based, prospective birth cohort study. *The Lancet* 2023; **401**. DOI:10.1016/S0140-6736(23)00811-5.

89 Coutts J, Fullarton J, Morris C, *et al.* Association between respiratory syncytial virus hospitalization in infancy and childhood asthma. *Pediatr Pulmonol* 2020; **55**. DOI:10.1002/ppul.24676.

90 Binns E, Tuckerman J, Licciardi P V., Wurzel D. Respiratory syncytial virus, recurrent wheeze and asthma: A narrative review of pathophysiology, prevention and future directions. J. Paediatr. Child Health. 2022; **58**. DOI:10.1111/jpc.16197.

91 Garegnani L, Styrmisdóttir L, Roson Rodriguez P, Escobar Liquitay CM, Esteban I, Franco JVA. Palivizumab for preventing severe respiratory syncytial virus (RSV) infection in children. Cochrane Database of Systematic Reviews. 2021; **2021**. DOI:10.1002/14651858.CD013757.pub2.

92 Mazur NI, Löwensteyn YN, Terstappen J, *et al.* Daily intranasal palivizumab to prevent respiratory syncytial virus infection in healthy preterm infants: a phase 1/2b randomized placebo-controlled trial. *EClinicalMedicine* 2023; **66**. DOI:10.1016/j.eclinm.2023.102324.

93 O'Brien KL, Chandran A, Weatherholtz R, *et al.* Efficacy of motavizumab for the prevention of respiratory syncytial virus disease in healthy Native American infants: A phase 3 randomised double-blind placebo-controlled trial. *Lancet Infect Dis* 2015; **15**. DOI:10.1016/S1473-3099(15)00247-9.

94 McLellan JS, Ray WC, Peeples ME. Structure and Function of RSV Surface Glycoproteins. *Curr Top Microbiol Immunol* 2013; **372**.

95 McLellan JS, Chen M, Kim A, Yang Y, Graham BS, Kwong PD. Structural basis of respiratory syncytial virus neutralization by motavizumab. *Nat Struct Mol Biol* 2010; **17**. DOI:10.1038/nsmb.1723.

96 Jha A, Jarvis H, Fraser C, Openshaw PJ. SARS, MERS and other Viral Lung Infections. In: Hui DS, Rossi GA, Johnston SL, eds. SARS, MERS and other Viral Lung Infections. European Respiratory Society, 2016. DOI:10.1183/2312508x.erm7216.

97 McLellan JS, Chen M, Joyce MG, *et al.* Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science (1979)* 2013; **342**. DOI:10.1126/science.1243283.

98 Leroux-Roels I, Davis MG, Steenackers K, *et al.* Safety and Immunogenicity of a Respiratory Syncytial Virus Prefusion F (RSVPreF3) Candidate Vaccine in Older Adults: Phase 1/2 Randomized Clinical Trial. *Journal of Infectious Diseases* 2023; **227**. DOI:10.1093/infdis/jiac327.

99 Papi A, Ison MG, Langley JM, *et al.* Respiratory Syncytial Virus Prefusion F Protein Vaccine in Older Adults. *New England Journal of Medicine* 2023; **388**. DOI:10.1056/nejmoa2209604.

100 Walsh EE, Pérez Marc G, Zareba AM, *et al.* Efficacy and Safety of a Bivalent RSV Prefusion F Vaccine in Older Adults. *New England Journal of Medicine* 2023; **388**. DOI:10.1056/nejmoa2213836.

101 Wang L, Wang Y, Ye D, Liu Q. Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *Int J Antimicrob Agents* 2020; **55**. DOI:10.1016/j.ijantimicag.2020.105948.

102 Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of SARS-CoV-2 entry into cells. Nat. Rev. Mol. Cell Biol. 2022; **23**. DOI:10.1038/s41580-021-00418-x.

103 Cilloniz C, Torres A, Luna CM, Hurtado JC, Marcos MÁ. Respiratory viruses: their importance and lessons learned from COVID-19. *European Respiratory Review* 2022; **31**. DOI:10.1183/16000617.0051-2022.

104 Nicola M, Alsafi Z, Sohrabi C, *et al.* The socio-economic implications of the coronavirus pandemic (COVID-19): A review. International Journal of Surgery. 2020; **78**. DOI:10.1016/j.ijsu.2020.04.018.

105 Azagew AW, Beko ZW, Ferede YM, Mekonnen HS, Abate HK, Mekonnen CK. Global prevalence of COVID-19-induced acute respiratory distress syndrome: systematic review and meta-analysis. *Syst Rev* 2023; **12**. DOI:10.1186/s13643-023-02377-0.

106 Haque SKM, Ashwaq O, Sarief A, Azad John Mohamed AK. A comprehensive review about SARS-CoV-2. Future Virol. 2020; **15**. DOI:10.2217/fvl-2020-0124.

107 Mandala VS, McKay MJ, Shcherbakov AA, Dregni AJ, Kolocouris A, Hong M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat Struct Mol Biol* 2020; **27**. DOI:10.1038/s41594-020-00536-8.

108 Zhang Z, Nomura N, Muramoto Y, *et al.* Structure of SARS-CoV-2 membrane protein essential for virus assembly. *Nat Commun* 2022; **13**. DOI:10.1038/s41467-022-32019-3.

109 Sofi MS, Hamid A, Bhat SU. SARS-CoV-2: A critical review of its history, pathogenesis, transmission, diagnosis and treatment. Biosaf. Health. 2020; **2**. DOI:10.1016/j.bsheal.2020.11.002.

110 Peng Y, Du N, Lei Y, *et al.* Structures of the SARS -CoV-2 nucleocapsid and their perspectives for drug design . *EMBO J* 2020; **39**. DOI:10.15252/embj.2020105938.

111 Chang CK, Jeyachandran S, Hu NJ, *et al.* Structure-based virtual screening and experimental validation of the discovery of inhibitors targeted towards the human coronavirus nucleocapsid protein. *Mol Biosyst* 2016; **12**. DOI:10.1039/c5mb00582e.

112 Amraei R, Yin W, Napoleon MA, *et al.* CD209L/L-SIGN and CD209/DC-SIGN Act as Receptors for SARS-CoV-2. *ACS Cent Sci* 2021; **7**. DOI:10.1021/acscentsci.0c01537.

113 Wang S, Qiu Z, Hou Y, *et al.* AXL is a candidate receptor for SARS-CoV-2 that promotes infection of pulmonary and bronchial epithelial cells. *Cell Res* 2021; **31**. DOI:10.1038/s41422-020-00460-y.

114 Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA - Journal of the American Medical Association. 2020; **324**. DOI:10.1001/jama.2020.12839.

115 Cascella M, Rajnik M, Aleem A, Dulebohn SC, Raffaela Di Napoli. Features, Evaluation, and Treatment of Coronavirus (COVID-19). *Natinal library of medicine* 2022; **2019**.

116 Vihta KD, Pouwels KB, Peto TEA, *et al.* Symptoms and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Positivity in the General Population in the United Kingdom. *Clinical Infectious Diseases* 2022; **75**. DOI:10.1093/cid/ciab945.

117 Xu J, Chu M, Zhong F, *et al.* Digestive symptoms of COVID-19 and expression of ACE2 in digestive tract organs. *Cell Death Discov* 2020; **6**. DOI:10.1038/s41420-020-00307-w.

118 Ousseiran ZH, Fares Y, Chamoun WT. Neurological manifestations of COVID-19: a systematic review and detailed comprehension. International Journal of Neuroscience. 2023; **133**. DOI:10.1080/00207454.2021.1973000.

119 Yoo J, Kim JH, Jeon J, Kim J, Song TJ. Risk of COVID-19 Infection and of Severe Complications Among People With Epilepsy: A Nationwide Cohort Study. *Neurology* 2022; **98**. DOI:10.1212/WNL.0000000000200195.

120 Desai AD, Lavelle M, Boursiquot BC, Wan EY. Long-term complications of COVID-19. Am. J. Physiol. Cell Physiol. 2022; **322**. DOI:10.1152/AJPCELL.00375.2021.

121 Michalski JE, Kurche JS, Schwartz DA. From ARDS to pulmonary fibrosis: the next phase of the COVID-19 pandemic? Translational Research. 2022; **241**. DOI:10.1016/j.trsl.2021.09.001.

122 Pfortmueller CA, Spinetti T, Urman RD, Luedi MM, Schefold JC. COVID-19-associated acute respiratory distress syndrome (CARDS): Current knowledge on pathophysiology and ICU treatment – A narrative review. Best Pract. Res. Clin. Anaesthesiol. 2021; **35**. DOI:10.1016/j.bpa.2020.12.011.

123 Carsana L, Sonzogni A, Nasr A, *et al.* Pulmonary post-mortem findings in a series of COVID-19 cases from northern Italy: a two-centre descriptive study. *Lancet Infect Dis* 2020; **20**. DOI:10.1016/S1473-3099(20)30434-5.

124 Kim Y, Bae S, Chang HH, Kim SW. Long COVID prevalence and impact on quality of life 2 years after acute COVID-19. *Sci Rep* 2023; **13**. DOI:10.1038/s41598-023-36995-4.

125 Gheorghita R, Soldanescu I, Lobiuc A, *et al.* The knowns and unknowns of long COVID-19: from mechanisms to therapeutical approaches. Front. Immunol. 2024; **15**. DOI:10.3389/fimmu.2024.1344086.

126 Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat. Rev. Microbiol. 2023; **21**. DOI:10.1038/s41579-022-00846-2.

127 Li G, Hilgenfeld R, Whitley R, De Clercq E. Therapeutic strategies for COVID-19: progress and lessons learned. Nat. Rev. Drug Discov. 2023; **22**. DOI:10.1038/s41573-023-00672-y.

128 Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 2014; **30**. DOI:10.1093/bioinformatics/btt703.

129 Graña C, Ghosn L, Evrenoglou T, *et al.* Efficacy and safety of COVID-19 vaccines. Cochrane Database of Systematic Reviews. 2022; **2022**. DOI:10.1002/14651858.CD015477.

130 Polack FP, Thomas SJ, Kitchin N, *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* 2020; **383**. DOI:10.1056/nejmoa2034577.

131 Baden LR, El Sahly HM, Essink B, *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine* 2021; **384**. DOI:10.1056/nejmoa2035389.

132 Voysey M, Clemens SAC, Madhi SA, *et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet* 2021; **397**. DOI:10.1016/S0140-6736(20)32661-1.

133 Heath PT, Galiza EP, Baxter DN, *et al.* Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine. *New England Journal of Medicine* 2021; **385**. DOI:10.1056/nejmoa2107659.

134 Thomsen SF, Stensballe LG, Skytthe A, Kyvik KO, Backer V, Bisgaard H. Increased concordance of severe respiratory syncytial virus infection in identical twins. *Pediatrics* 2008; **121**. DOI:10.1542/peds.2007-1889.

135 Nogales A, Dediego ML. Host single nucleotide polymorphisms modulating influenza a virus disease in humans. Pathogens. 2019; **8**. DOI:10.3390/pathogens8040168.

136 Peterson RE, Kuchenbaecker K, Walters RK, *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. Cell. 2019; **179**. DOI:10.1016/j.cell.2019.08.051.

137 Jones SC, Cardone KM, Bradford Y, Tishkoff SA, Ritchie MD. The Impact of Ancestry on Genome-Wide Association Studies. *Pac Symp Biocomput* 2025; **30**. DOI:10.1142/9789819807024_0019.

138 Tanaka T, Narazaki M, Kishimoto T. Il-6 in inflammation, Immunity, And disease. *Cold Spring Harb Perspect Biol* 2014; **6**. DOI:10.1101/cshperspect.a016295.

139 Edwards MR, Haas J, Panettieri RA, Johnson M, Johnston SL. Corticosteroids and β2 agonists differentially regulate rhinovirus-induced interleukin-6 via distinct cis-acting elements. *Journal of Biological Chemistry* 2007; **282**. DOI:10.1074/jbc.M701325200.

140 Yamamoto F, Yamamoto M, Muñiz-Diaz E. Blood group ABO polymorphism inhibits SARS-CoV-2 infection and affects COVID-19 progression. Vox Sang. 2021; **116**. DOI:10.1111/vox.13004.

141 Terry CF, Loukaci V, Green FR. Cooperative influence of genetic polymorphisms on interleukin 6 transcriptional regulation. *Journal of Biological Chemistry* 2000; **275**. DOI:10.1074/jbc.M000379200.

142 Doyle WJ, Casselbrant ML, Li-Korotky HS, *et al.* The interleukin 6-174 C/C genotype predicts greater rhinovirus illness. *Journal of Infectious Diseases* 2010; **201**. DOI:10.1086/649559.

143 Sabat R, Grütz G, Warszawska K, *et al.* Biology of interleukin-10. *Cytokine Growth Factor Rev* 2010; **21**. DOI:10.1016/j.cytogfr.2010.09.002.

144 Rowe RK, Pyle DM, Farrar JD, Gill MA. IgE-mediated regulation of IL-10 and type I IFN enhances rhinovirus-induced Th2 differentiation by primary human monocytes. *Eur J Immunol* 2020; **50**. DOI:10.1002/eji.201948396.

145 Gao J, Wei L, Fu R, *et al.* Association of interleukin-10 polymorphisms (rs1800872, rs1800871, and rs1800896) with predisposition to IgA nephropathy in a Chinese han population: A case-control study. *Kidney Blood Press Res* 2017; **42**. DOI:10.1159/000471899.

146 Holster A, Teräsjärvi J, Vuononvirta J, *et al.* Polymorphisms in the promoter region of IL10 gene are associated with virus etiology of infant bronchiolitis. *World Journal of Pediatrics* 2018; **14**. DOI:10.1007/s12519-018-0161-7.

147 Helminen M, Nuolivirta K, Virta M, *et al.* IL-10 gene polymorphism at -1082 A/G is associated with severe rhinovirus bronchiolitis in infants. *Pediatr Pulmonol* 2008; **43**. DOI:10.1002/ppul.20793.

148 Antosova M, Mokra D, Pepucha L, *et al.* Physiology of nitric oxide in the respiratory system. Physiol. Res. 2017; **66**. DOI:10.33549/physiolres.933673.

149 Sanders SP, Proud D, Permutt S, Siekierski ES, Yachechko R, Liu MC. Role of nasal nitric oxide in the resolution of experimental rhinovirus infection. *Journal of Allergy and Clinical Immunology* 2004; **113**. DOI:10.1016/j.jaci.2004.01.755.

150 Paul KC, Sinsheimer JS, Rhodes SL, Cockburn M, Bronstein J, Ritz B. Organophosphate pesticide exposures, nitric oxide synthase gene variants, and gene–pesticide interactions in a case–control study of parkinson's disease, California (USA). *Environ Health Perspect* 2016; **124**. DOI:10.1289/ehp.1408976.

151 Alvarez AE, Marson FAL, Bertuzzo CS, *et al.* Association between single nucleotide polymorphisms in TLR4, TLR2, TLR9, VDR, NOS2 and CCL5 genes with acute viral bronchiolitis. *Gene* 2018; **645**. DOI:10.1016/j.gene.2017.12.022.

152 Holgate ST, Yang Y, Haitchi HM, *et al.* The genetics of asthma: ADAM33 as an example of a susceptibility gene. In: Proceedings of the American Thoracic Society. 2006. DOI:10.1513/pats.200603-026AW.

153 Tripathi P, Awasthi S, Gao P. ADAM metallopeptidase domain 33 (ADAM33): A promising target for asthma. Mediators Inflamm. 2014; **2014**. DOI:10.1155/2014/572025.

154 Drysdale SB, Prendergast M, Alcazar M, *et al.* Genetic predisposition of RSV infection-related respiratory morbidity in preterm infants. *Eur J Pediatr* 2014; **173**. DOI:10.1007/s00431-014-2263-0.

155 Appay V, Rowland-Jones SL. RANTES: A versatile and controversial chemokine. Trends Immunol. 2001; **22**. DOI:10.1016/S1471-4906(00)01812-3.

156 Marques RE, Guabiraba R, Russo RC, Teixeira MM. Targeting CCL5 in inflammation. Expert Opin. Ther. Targets. 2013; **17**. DOI:10.1517/14728222.2013.837886.

157 Culley FJ, Pennycook AMJ, Tregoning JS, *et al.* Role of CCL5 (RANTES) in Viral Lung Disease. *J Virol* 2006; **80**. DOI:10.1128/jvi.00496-06.

158 An P, Nelson GW, Wang L, *et al.* Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc Natl Acad Sci U S A* 2002; **99**. DOI:10.1073/pnas.142313799.

159 Hattori S, Shimojo N, Mashimo Y, *et al.* Relationship between RANTES polymorphisms and respiratory syncytial virus bronchiolitis in a Japanese infant population. *Jpn J Infect Dis* 2011; **64**.

160 Oppermann M. Chemokine receptor CCR5: Insights into structure, function, and regulation. Cell. Signal. 2004; **16**. DOI:10.1016/j.cellsig.2004.04.007.

161 Gomulska M, Rusin G, Gwiazdak P. Prevalence of CCR5-delta32 mutation in asthmatic and non-asthmatic subjects from department of medicine, JUCM, Cracow. *Folia Med Cracov* 2014; **54**.

162 Hull J, Rowlands K, Lockhart E, Moore C, Sharland M, Kwiatkowski D. Variants of the chemokine receptor CCR5 are associated with severe bronchiolitis caused by respiratory syncytial virus. *Journal of Infectious Diseases* 2003; **188**. DOI:10.1086/377587.

163 Sharygin D, Koniaris LG, Wells C, Zimmers TA, Hamidi T. Role of CD14 in human disease. Immunology. 2023; **169**. DOI:10.1111/imm.13634.

164 Kurt-Jones EA, Popova L, Kwinn L, *et al.* Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nat Immunol* 2000; **1**. DOI:10.1038/80833.

165 Ferrero E, Hsieh CL, Francke U, Goyert SM. CD14 is a member of the family of leucine-rich proteins and is encoded by a gene syntenic with multiple receptor genes. *The Journal of Immunology* 1990; **145**. DOI:10.4049/jimmunol.145.1.331.

166 Inoue Y, Shimojo N, Suzuki Y, *et al.* CD14 -550 C/T, which is related to the serum level of soluble CD14, is associated with the development of respiratory syncytial virus bronchiolitis in the Japanese population. *Journal of Infectious Diseases* 2007; **195**. DOI:10.1086/516790.

167 Liu QR, Pan CH, Hishimoto A, *et al.* Species differences in cannabinoid receptor 2 (CNR2 gene): Identification of novel human and rodent CB2 isoforms, differential tissue expression and regulation by cannabinoid receptor ligands. *Genes Brain Behav* 2009; **8**. DOI:10.1111/j.1601-183X.2009.00498.x.

168 Divín D, Goméz Samblas M, Kuttiyarthu Veetil N, *et al.* Cannabinoid receptor 2 evolutionary gene loss makes parrots more susceptible to neuroinflammation. *Proceedings of the Royal Society B: Biological Sciences* 2022; **289**. DOI:10.1098/rspb.2022.1941.

169 Zhang C, Ma J, Chen G, Fu D, Li L, Li M. Evaluation of common variants in CNR2 gene for bone mineral density and osteoporosis susceptibility in postmenopausal women of Han Chinese. *Osteoporosis International* 2015; **26**. DOI:10.1007/s00198-015-3195-x.

170 Tahamtan A, Samieipoor Y, Nayeri FS, *et al.* Effects of cannabinoid receptor type 2 in respiratory syncytial virus infection in human subjects and mice. *Virulence* 2018; **9**. DOI:10.1080/21505594.2017.1389369.

171 Anderson CS, Chu CY, Wang Q, *et al.* CX3CR1 as a respiratory syncytial virus receptor in pediatric human lung. *Pediatr Res* 2020; **87**. DOI:10.1038/s41390-019-0677-0.

172 Lee M, Lee Y, Song J, Lee J, Chang SY. Tissue-specific role of CX3CR1 expressing immune cells and their relationships with human disease. Immune Netw. 2018; **18**. DOI:10.4110/in.2018.18.e5.

173 Amanatidou V, Sourvinos G, Apostolakis S, Tsilimigaki A, Spandidos DA. T280M variation of the CX3C receptor gene is associated with increased risk for severe respiratory syncytial virus bronchiolitis. *Pediatric Infectious Disease Journal* 2006; **25**. DOI:10.1097/01.inf.0000214998.16248.b7.

174 Mosaad YM. Clinical Role of Human Leukocyte Antigen in Health and Disease. Scand. J. Immunol. 2015; **82**. DOI:10.1111/sji.12329.

175 Tumer G, Simpson B, Roberts TK. Genetics, Human Major Histocompatibility Complex (MHC). 2019.

176 Williams TM. Human leukocyte antigen gene polymorphism and the histocompatibility laboratory. Journal of Molecular Diagnostics. 2001; **3**. DOI:10.1016/S1525-1578(10)60658-7.

177 Salas A, Pardo-Seco J, Cebey-López M, *et al.* Whole Exome Sequencing reveals new candidate genes in host genomic susceptibility to Respiratory Syncytial Virus Disease. *Sci Rep* 2017; **7**. DOI:10.1038/s41598-017-15752-4.

178 Kak G, Raza M, Tiwari BK. Interferon-gamma (IFN-γ): Exploring its implications in infectious diseases. Biomol. Concepts. 2018; **9**. DOI:10.1515/bmc-2018-0007.

179 Gentile DA, Doyle WJ, Zeevi A, *et al.* Cytokine gene polymorphisms moderate illness severity in infants with respiratory syncytial virus infection. *Hum Immunol* 2003; **64**. DOI:10.1016/S0198-8859(02)00827-3.

180 Huang J, Zhang M, Zhang X, Lu A, Wang L, Chen C. IFN-γ CA microsatellite polymorphism is associated with susceptibility to respiratory syncytial virus Infection and severity. *Acta Paediatrica, International Journal of Paediatrics* 2014; **103**. DOI:10.1111/apa.12767.

181 Haas H, Falcone FH, Holland MJ, *et al.* Early interleukin-4: Its role in the switch towards a Th2 response and IgE-mediated allergy. Int. Arch. Allergy Immunol. 1999; **119**. DOI:10.1159/000024182.

182 Legg JP, Hussain IR, Warner JA, Johnston SL, Warner JO. Type 1 and type 2 cytokine imbalance in acute respiratory syncytial virus bronchiolitis. *Am J Respir Crit Care Med* 2003; **168**. DOI:10.1164/rccm.200210-1148OC.

183 Kruse S, Braun S, Deichmann KA. Distinct signal transduction processes by IL-4 and IL-13 and influences from the Q551R variant of the human IL-4 receptor alpha chain. *Respir Res* 2002; **3**. DOI:10.1186/rr174.

184 Kousha A, Mahdavi Gorabi A, Forouzesh M, *et al.* Interleukin 4 gene polymorphism (−589C/T) and the risk of asthma: a meta-analysis and met-regression based on 55 studies. *BMC Immunol* 2020; **21**. DOI:10.1186/s12865-020-00384-7.

185 Hoebee B, Rietveld E, Bont L, *et al.* Association of severe respiratory syncytial virus bronchiolitis with interleukin-4 and interleukin-4 receptor α polymorphisms. *Journal of Infectious Diseases* 2003; **187**. DOI:10.1086/345859.

186 Zhang M, Lu Y, Zhang X, Lu A, Wang L, Chen C. Interleukin-4 polymorphism is associated with severity of respiratory syncytial virus infection. *J Paediatr Child Health* 2016; **52**. DOI:10.1111/jpc.12985.

187 Forton JT, Rowlands K, Rockett K, *et al.* Genetic association study for RSV bronchiolitis in infancy at the 5q31 cytokine cluster. *Thorax* 2009; **64**. DOI:10.1136/thx.2008.102111.

188 Marr N, Hirschfeld AF, Lam A, Wang S, Lavoie PM, Turvey SE. Assessment of genetic associations between common single nucleotide polymorphisms in RIG-I-like receptor and IL-4 signaling genes and severe respiratory syncytial virus infection in children: A candidate gene case-control study. *PLoS One* 2014; **9**. DOI:10.1371/journal.pone.0100269.

189 Puthothu B, Krueger M, Forster J, Heinzmann A. Association between severe respiratory syncytial virus infection and IL13/IL4 haplotypes. *Journal of Infectious Diseases* 2006; **193**. DOI:10.1086/499316.

190 Tapia LI, Ampuero S, Palomino MA, *et al.* Respiratory syncytial virus infection and recurrent wheezing in Chilean infants: A genetic background? *Infection, Genetics and Evolution* 2013; **16**. DOI:10.1016/j.meegid.2012.12.030.

191 Pyle CJ, Uwadiae FI, Swieboda DP, Harker JA. Early IL-6 signalling promotes IL-27 dependent maturation of regulatory T cells in the lungs and resolution of viral immunopathology. *PLoS Pathog* 2017; **13**. DOI:10.1371/journal.ppat.1006640.

192 Vázquez Y, González L, Noguera L, *et al.* Cytokines in the respiratory airway as biomarkers of severity and prognosis for respiratory syncytial virus infection: An update. Front. Immunol. 2019; **10**. DOI:10.3389/fimmu.2019.01154.

193 Matsushima K, Oppenheim JJ. Interleukin 8 and MCAF: novel inflammatory cytokines inducible by IL 1 and TNF. Cytokine. 1989; **1**. DOI:10.1016/1043-4666(89)91043-0.

194 Hull J, Thomson A, Kwiatkowski D. Association of respiratory syncytial virus bronchiolitis with the interleukin 8 gene region in UK families. *Thorax* 2000; **55**. DOI:10.1136/thorax.55.12.1023.

195 Goetghebuer T, Isles K, Moore C, Thomson A, Kwiatkowski D, Hull J. Genetic predisposition to wheeze following respiratory syncytial virus bronchiolitis. *Clinical and Experimental Allergy* 2004; **34**. DOI:10.1111/j.1365-2222.2004.1947.x.

196 Lu A, Wang L, Zhang X. Haplotype of IL-8 -251T and 781C is associated with the susceptibility to respiratory syncytial virus. *J Trop Pediatr* 2009; **56**. DOI:10.1093/tropej/fmp101.

197 Iyer SS, Cheng G. Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease. *Crit Rev Immunol* 2012; **32**. DOI:10.1615/critrevimmunol.v32.i1.30.

198 Panuska JR, Merolla R, Rebert NA, *et al.* Respiratory syncytial virus induces interleukin-10 by human alveolar macrophages: Suppression of early cytokine production and implications for incomplete immunity. *Journal of Clinical Investigation* 1995; **96**. DOI:10.1172/jci118302.

199 Khurana Hershey GK. IL-13 receptors and signaling pathways: An evolving web. Journal of Allergy and Clinical Immunology. 2003; **111**. DOI:10.1067/mai.2003.1333.

200 Junttila IS. Tuning the cytokine responses: An update on interleukin (IL)-4 and IL-13 receptor complexes. Front. Immunol. 2018; **9**. DOI:10.3389/fimmu.2018.00888.

201 Rael EL, Lockey RF. Interleukin-13 signaling and its role in asthma. World Allergy Organization Journal. 2011; **4**. DOI:10.1097/WOX.0b013e31821188e0.

202 Ermers MJJ, Hoebee B, Hodemaekers HM, Kimman TG, Kimpen JLL, Bont L. IL-13 genetic polymorphism identifies children with late wheezing after respiratory syncytial virus infection. *Journal of Allergy and Clinical Immunology* 2007; **119**. DOI:10.1016/j.jaci.2006.12.655.

203 Choi EH, Lee HJ, Yoo T, Chanock SJ. A common haplotype of interleukin-4 gene il4 is associated with severe respiratory syncytial virus disease in korean children. *Journal of Infectious Diseases* 2002; **186**. DOI:10.1086/344310.

204 Perera PY, Lichy JH, Waldmann TA, Perera LP. The role of interleukin-15 in inflammation and immune responses to infection: Implications for its therapeutic use. Microbes Infect. 2012; **14**. DOI:10.1016/j.micinf.2011.10.006.

205 Zdrenghea MT, Telcian AG, Laza-Stanca V, *et al.* RSV infection modulates IL-15 production and MICA levels in respiratory epithelial cells. *European Respiratory Journal* 2012; **39**. DOI:10.1183/09031936.00099811.

206 Pagliari D, Cianci R, Frosali S, *et al.* The role of IL-15 in gastrointestinal diseases: A bridge between innate and adaptive immune response. Cytokine Growth Factor Rev. 2013; **24**. DOI:10.1016/j.cytogfr.2013.05.004.

207 Yamamoto-Furusho JK, De-León-Rendón JL, Álvarez-León E, Valdivia-Eziquio M, Obil-Chavarría C, Vargas-Alarcón G. Association of the interleukin 15 (IL-15) gene polymorphisms with the risk of developing ulcerative colitis in Mexican individuals. *Mol Biol Rep* 2014; **41**. DOI:10.1007/s11033-014-3067-z.

208 Janssen R, Bont L, Siezen CLE, *et al.* Genetic susceptibility to respiratory syncytial virus bronchiolitis is predominantly associated with innate immune genes. *Journal of Infectious Diseases* 2007; **196**. DOI:10.1086/520886.

209 Xu S, Cao X. Interleukin-17 and its expanding biological functions. Cell. Mol. Immunol. 2010; **7**. DOI:10.1038/cmi.2010.21.

210 Long X, Xie J, Ren L, *et al.* IL-17A plays a critical role in RSV infection in children and mice. *Virol J* 2023; **20**. DOI:10.1186/s12985-023-01990-8.

211 Gaffen SL. Structure and signalling in the IL-17 receptor family. Nat. Rev. Immunol. 2009; **9**. DOI:10.1038/nri2586.

212 Harker JA, Godlee A, Wahlsten JL, *et al.* Interleukin 18 Coexpression during Respiratory Syncytial Virus Infection Results in Enhanced Disease Mediated by Natural Killer Cells. *J Virol* 2010; **84**. DOI:10.1128/jvi.02014-09.

213 Van Benten IJ, Van Drunen CM, Koopman LP, *et al.* RSV-induced bronchiolitis but not upper respiratory tract infection is accompanied by an increased nasal IL-18 response. *J Med Virol* 2003; **71**. DOI:10.1002/jmv.10482.

214 Puthothu B, Krueger M, Forster J, Heinze J, Weckmann M, Heinzmann A. Interleukin (IL)-18 polymorphism 133C/G is associated with severe respiratory syncytial virus infection. *Pediatric Infectious Disease Journal* 2007; **26**. DOI:10.1097/INF.0b013e3181453579.

215 Kinkel SA, Galeev R, Flensburg C, *et al.* Jarid2 regulates hematopoietic stem cell function by acting with polycomb repressive complex 2. *Blood* 2015; **125**. DOI:10.1182/blood-2014-10-603969.

216 Bergé-Lefranc JL, Jay P, Massacrier A, *et al.* Characterization of the human jumonji gene. *Hum Mol Genet* 1996; **5**. DOI:10.1093/hmg/5.10.1637.

217 Potapova O, Basu S, Mercola D, Holbrook NJ. Protective Role for c-Jun in the Cellular Response to DNA Damage. *Journal of Biological Chemistry* 2001; **276**. DOI:10.1074/jbc.M102075200.

218 Gustems M, Woellmer A, Rothbauer U, *et al.* C-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. *Nucleic Acids Res* 2014; **42**. DOI:10.1093/nar/gkt1323.

219 Xing Q, Feng Y, Sun H, *et al.* Scavenger receptor MARCO contributes to macrophage phagocytosis and clearance of tumor cells. *Exp Cell Res* 2021; **408**. DOI:10.1016/j.yexcr.2021.112862.

220 Elshourbagy NA, Li X, Terrett J, *et al.* Molecular characterization of a human scavenger receptor, human MARCO. *Eur J Biochem* 2000; **267**. DOI:10.1046/j.1432-1327.2000.01077.x.

221 High M, Cho HY, Marzec J, *et al.* Determinants of host susceptibility to murine respiratory syncytial virus (RSV) disease identify a role for the innate immunity scavenger receptor MARCO gene in human infants. *EBioMedicine* 2016; **11**. DOI:10.1016/j.ebiom.2016.08.011.

222 Roth L, Kalev-Altman R, Monsonego-Ornan E, Sela-Donenfeld D. A new role of the membrane-type matrix metalloproteinase 16 (MMP16/MT3-MMP) in neural crest cell migration. *International Journal of Developmental Biology* 2017; **61**. DOI:10.1387/ijdb.160286ds.

223 Verma RP, Hansch C. Matrix metalloproteinases (MMPs): Chemical-biological functions and (Q)SARs. Bioorg. Med. Chem. 2007; **15**. DOI:10.1016/j.bmc.2007.01.011.

224 Khalil RA. Matrix Metalloproteinases and Tissue Remodeling in Health and Disease: Cardiovascular Remodeling. 2017.

225 Hadchouel A, Decobert F, Franco-Montoya ML, *et al.* Matrix metalloproteinase gene polymorphisms and bronchopulmonary dysplasia: Identification of MMP16 as a new player in lung development. *PLoS One* 2008; **3**. DOI:10.1371/journal.pone.0003188.

226 Vasseur R, Skrypek N, Duchêne B, *et al.* The mucin MUC4 is a transcriptional and post-transcriptional target of K-ras oncogene in pancreatic cancer. Implication of MAPK/AP-1, NF-κB and RalB signaling pathways. *Biochim Biophys Acta Gene Regul Mech* 2015; **1849**. DOI:10.1016/j.bbagrm.2015.10.014.

227 Chakraborty S, Jain M, Sasson AR, Batra SK. MUC4 as a diagnostic marker in cancer. Expert Opin. Med. Diagn. 2008; **2**. DOI:10.1517/17530059.2.8.891.

228 Schwab LSU, Villalón-Letelier F, Tessema MB, *et al.* Expression of a Functional Mx1 Protein Is Essential for the Ability of RIG-I Agonist Prophylaxis to Provide Potent and Long-Lasting Protection in a Mouse Model of Influenza A Virus Infection. *Viruses* 2022; **14**. DOI:10.3390/v14071547.

229 Verhelst J, Hulpiau P, Saelens X. Mx Proteins: Antiviral Gatekeepers That Restrain the Uninvited. *Microbiology and Molecular Biology Reviews* 2013; **77**. DOI:10.1128/mmbr.00024-13.

230 Ciencewicki JM, Wang X, Marzec J, *et al.* A genetic model of differential susceptibility to human respiratory syncytial virus (RSV) infection. *FASEB Journal* 2014; **28**. DOI:10.1096/fj.13-239855.

231 Qin K, Zheng Z, He Y, *et al.* High expression of neutrophil cytosolic factor 2 (NCF2) is associated with aggressive features and poor prognosis of esophageal squamous cell carcinoma. *Int J Clin Exp Pathol* 2020; **13**.

232 Huang N, Zhang J, Kuang S, *et al.* Role of NCF2 as a potential prognostic factor and immune infiltration indicator in hepatocellular carcinoma. *Cancer Med* 2023; **12**. DOI:10.1002/cam4.5597.

233 Crist RC, Berrettini WH. Pharmacogenetics of OPRM1. Pharmacol. Biochem. Behav. 2014; **123**. DOI:10.1016/j.pbb.2013.10.018.

234 Yu Z, Wen L, Shen X, Zhang H. Effects of the OPRM1 A118G Polymorphism (rs1799971) on Opioid Analgesia in Cancer Pain: A Systematic Review and Meta-Analysis. *Clin J Pain* 2019; **35**.

235 Salimi V, Hennus MP, Mokhtari-Azad T, *et al.* Opioid receptors control viral replication in the airways. *Crit Care Med* 2013; **41**. DOI:10.1097/CCM.0b013e31826767a8.

236 Glezer I, Malnic B. Olfactory receptor function. In: Handbook of Clinical Neurology. 2019. DOI:10.1016/B978-0-444-63855-7.00005-8.

237 Weber L, Maßberg D, Becker C, *et al.* Olfactory receptors as biomarkers in human breast carcinoma tissues. *Front Oncol* 2018; **8**. DOI:10.3389/fonc.2018.00033.

238 Riel D Van, Verdijk R, Kuiken T. The olfactory nerve: A shortcut for influenza and other viral diseases into the central nervous system. *Journal of Pathology* 2015; **235**. DOI:10.1002/path.4461.

239 Li Y, Guo D, Ren M, *et al.* Long non-coding RNA SNAI3-AS1 promotes the proliferation and metastasis of hepatocellular carcinoma by regulating the UPF1/Smad7 signalling pathway. *J Cell Mol Med* 2019; **23**. DOI:10.1111/jcmm.14513.

240 Vieira F, Kung JW, Bhatti F. Structure, genetics and function of the pulmonary associated surfactant proteins A and D: The extra-pulmonary role of these C type lectins. Annals of Anatomy. 2017; **211**. DOI:10.1016/j.aanat.2017.03.002.

241 Nogee LM. Genetic causes of surfactant protein abnormalities. Curr. Opin. Pediatr. 2019; **31**. DOI:10.1097/MOP.0000000000000751.

242 Löfgren J, Rämet M, Renko M, Marttila R, Hallman M. Association between surfactant protein A gene locus and severe respiratory syncytial virus infection in infants. *Journal of Infectious Diseases* 2002; **185**. DOI:10.1086/338473.

243 Ampuero S, Luchsinger V, Tapia L, Palomino MA, Larrañaga CE. SP-A1, SP-A2 and SP-D gene polymorphisms in severe acute respiratory syncytial infection in Chilean infants. *Infection, Genetics and Evolution* 2011; **11**. DOI:10.1016/j.meegid.2011.04.033.

244 Lahti M, Löfgren J, Marttila R, *et al.* Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 2002; **51**. DOI:10.1203/00006450-200206000-00006.

245 Chu WM. Tumor necrosis factor. Cancer Lett. 2013; **328**: 222–5.

246 Bradley JR. TNF-mediated inflammatory disease. Journal of Pathology. 2008; **214**. DOI:10.1002/path.2287.

247 Santos LD, Antunes KH, Muraro SP, *et al.* TNF-mediated alveolar macrophage necroptosis drives disease pathogenesis during respiratory syncytial virus infection. *European Respiratory Journal* 2021; **57**. DOI:10.1183/13993003.03764-2020.

248 Kong DH, Kim YK, Kim MR, Jang JH, Lee S. Emerging roles of vascular cell adhesion molecule-1 (VCAM-1) in immunological disorders and cancer. Int. J. Mol. Sci. 2018; **19**. DOI:10.3390/ijms19041057.

249 Wang SZ, Hallsworth PG, Dowling KD, Alpers JH, Bowden JJ, Forsyth KD. Adhesion molecule expression on epithelial cells infected with respiratory syncytial virus. *European Respiratory Journal* 2000; **15**. DOI:10.1034/j.1399-3003.2000.15b23.x.

250 Kawasaki T, Kawai T. Toll-like receptor signaling pathways. Front. Immunol. 2014; **5**. DOI:10.3389/fimmu.2014.00461.

251 Molteni M, Gemma S, Rossetti C. The Role of Toll-Like Receptor 4 in Infectious and Noninfectious Inflammation. Mediators Inflamm. 2016; **2016**. DOI:10.1155/2016/6978936.

252 Tal G, Mandelberg A, Dalal I, *et al.* Association between common toll-like receptor 4 mutations and severe respiratory syncytial virus disease. *Journal of Infectious Diseases* 2004; **189**. DOI:10.1086/420830.

253 Awomoyi AA, Rallabhandi P, Pollin TI, *et al.* Association of TLR4 Polymorphisms with Symptomatic Respiratory Syncytial Virus Infection

in High-Risk Infants and Young Children. *The Journal of Immunology* 2007; **179**. DOI:10.4049/jimmunol.179.5.3171.

254 Paulus SC, Hirschfeld AF, Victor RE, Brunstein J, Thomas E, Turvey SE. Common human Toll-like receptor 4 polymorphisms-Role in susceptibility to respiratory syncytial virus infection and functional immunological relevance. *Clinical Immunology* 2007; **123**. DOI:10.1016/j.clim.2007.03.003.

255 Moen SH, Ehrnström B, Kojen JF, *et al.* Human Toll-like receptor 8 (TLR8) is an important sensor of pyogenic bacteria, and is attenuated by cell surface TLR signaling. *Front Immunol* 2019; **10**. DOI:10.3389/fimmu.2019.01209.

256 Youness A, Cenac C, Faz-López B, *et al.* TLR8 escapes X chromosome inactivation in human monocytes and CD4+ T cells. *Biol Sex Differ* 2023; **14**. DOI:10.1186/s13293-023-00544-5.

257 Siddiqui M, Manansala JS, Abdulrahman HA, *et al.* Immune modulatory effects of vitamin d on viral infections. Nutrients. 2020; **12**. DOI:10.3390/nu12092879.

258 Hansdottir S, Monick MM, Lovan N, Powers L, Gerke A, Hunninghake GW. Vitamin D Decreases Respiratory Syncytial Virus Induction of NF-κB–Linked Chemokines and Cytokines in Airway Epithelium While Maintaining the Antiviral State. *The Journal of Immunology* 2010; **184**. DOI:10.4049/jimmunol.0902840.

259 Gross C, Eccleshall TR, Malloy PJ, Villa ML, Marcus R, Feldman D. The presence of a polymorphism at the translation initiation site of the vitamin D receptor gene is associated with low bone mineral density in postmenopausal Mexican-American women. *Journal of Bone and Mineral Research* 1996; **11**. DOI:10.1002/jbmr.5650111204.

260 Kresfelder TL, Janssen R, Bont L, Venter M. Confirmation of an association between single nucleotide polymorphisms in the VDR gene with respiratory syncytial virus related disease in South African Children. *J Med Virol* 2011; **83**. DOI:10.1002/jmv.22179.

261 Hosseini-Maaf B, Hellberg Å, Rodrigues MJ, Chester MA, Olsson ML. ABO exon and intron analysis in individuals with the AweakB phenotype reveals a novel Olv-A2 hybrid allele that causes four missense mutations in the A transferase. *BMC Genet* 2003; **4**. DOI:10.1186/1471-2156-4-17.

262 Wu Y, Feng Z, Li P, Yu Q. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. *Clinica Chimica Acta* 2020; **509**. DOI:10.1016/j.cca.2020.06.026.

263 Hoiland RL, Fergusson NA, Mitra AR, *et al.* The association of ABO blood group with indices of disease severity and multiorgan dysfunction in COVID-19. *Blood Adv* 2020; **4**. DOI:10.1182/BLOODADVANCES.2020002623.

264 D E, F D, L B, *et al.* Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *New England Journal of Medicine* 2020; **383**.

265 Niemi MEK, Karjalainen J, Liao RG, *et al.* Mapping the human genetic architecture of COVID-19. *Nature* 2021; **600**: 472–7.

266 Paterson AD, Lopes-Virella MF, Waggott D, *et al.* Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. *Arterioscler Thromb Vasc Biol* 2009; **29**. DOI:10.1161/ATVBAHA.109.192971.

267 Coates D. The angiotensin converting enzyme (ACE). International Journal of Biochemistry and Cell Biology. 2003; **35**. DOI:10.1016/S1357-2725(02)00309-6.

268 Khurana V, Goswami B. Angiotensin converting enzyme (ACE). Clinica Chimica Acta. 2022; **524**. DOI:10.1016/j.cca.2021.10.029.

269 Zmorzynski S, Szudy-Szczyrek A, Popek-Marciniec S, *et al.* ACEinsertion/deletion polymorphism (rs4646994) is associated with the increased risk of multiple myeloma. *Front Oncol* 2019; **9**. DOI:10.3389/fonc.2019.00044.

270 Hubacek JA, Dusek L, Majek O, *et al.* ACE I/D polymorphism in Czech first-wave SARS-CoV-2-positive survivors. *Clinica Chimica Acta* 2021; **519**. DOI:10.1016/j.cca.2021.04.024.

271 Verma S, Abbas M, Verma S, *et al.* Impact of I/D polymorphism of angiotensin-converting enzyme 1 (ACE1) gene on the severity of COVID-19 patients. *Infection, Genetics and Evolution* 2021; **91**. DOI:10.1016/j.meegid.2021.104801.

272 Bourgonje AR, Abdulle AE, Timens W, *et al.* Angiotensin-converting enzyme 2 (ACE2), SARS-CoV-2 and the pathophysiology of coronavirus disease 2019 (COVID-19). Journal of Pathology. 2020; **251**. DOI:10.1002/path.5471.

273 Gemmati D, Bramanti B, Serino ML, Secchiero P, Zauli G, Tisato V. COVID-19 and individual genetic susceptibility/receptivity: Role of ACE1/ACE2 genes, immunity, inflammation and coagulation. might the double x-chromosome in females be protective against SARS-COV-2 compared to the single x-chromosome in males? *Int J Mol Sci* 2020; **21**. DOI:10.3390/ijms21103474.

274 Möhlendick B, Schönfelder K, Breuckmann K, *et al.* ACE2 polymorphism and susceptibility for SARS-CoV-2 infection and severity of COVID-19. *Pharmacogenet Genomics* 2021. DOI:10.1097/FPC.0000000000000436.

275 Horowitz JE, Kosmicki JA, Damask A, *et al.* Genome-wide analysis provides genetic evidence that ACE2 influences COVID-19 risk and yields risk scores associated with severe disease. *Nat Genet* 2022; **54**. DOI:10.1038/s41588-021-01006-7.

276 Kurotaki A, Kuwata H, Hara S. Substrate Specificity of Human Long-Chain Acyl-CoA Synthetase ACSL6 Variants. *Biol Pharm Bull* 2021; **44**. DOI:10.1248/bpb.b21-00551.

277 Kang HJ, Lee SK, Kim MH, Choi HJ, Lee SH, Kwack KB. Acyl-CoA synthetase long-chain family member 6 is associated with premature

ovarian failure. *Fertil Steril* 2009; **91**. DOI:10.1016/j.fertnstert.2008.03.035.

278 Kousathanas A, Pairo-Castineira E, Rawlik K, *et al.* Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* 2022; **607**. DOI:10.1038/s41586-022-04576-6.

279 Dominiczak MH, Caslake MJ. Apolipoproteins: Metabolic role and clinical biochemistry applications. Ann. Clin. Biochem. 2011; **48**. DOI:10.1258/acb.2011.011111.

280 Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein e and Alzheimer disease: Risk, mechanisms and therapy. Nat. Rev. Neurol. 2013; **9**. DOI:10.1038/nrneurol.2012.263.

281 Zhang H, Shao L, Lin Z, *et al.* APOE interacts with ACE2 inhibiting SARS-CoV-2 cellular entry and inflammation in COVID-19 patients. *Signal Transduct Target Ther* 2022; **7**. DOI:10.1038/s41392-022-01118-4.

282 Tejedor MT, Garcia-Sobreviela MP, Ledesma M, Arbones-Mainar JM. The apolipoprotein e polymorphism rs7412 associates with body fatness independently of plasma lipids in middle aged men. *PLoS One* 2014; **9**. DOI:10.1371/journal.pone.0108605.

283 Espinosa-Salinas I, Colmenarejo G, Fernández-Díaz CM, *et al.* Potential protective effect against SARS-CoV-2 infection by APOE rs7412 polymorphism. *Sci Rep* 2022; **12**. DOI:10.1038/s41598-022-10923-4.

284 Hawkey-Noble A, Umali J, Fowler G, French CR. Expression of three P4-phospholipid flippases—atp11a, atp11b, and atp11c in zebrafish (Danio rerio). *Gene Expression Patterns* 2020; **36**. DOI:10.1016/j.gep.2020.119115.

285 Yin J, Xie X, Ye Y, Wang L, Che F. BCL11A: A potential diagnostic biomarker and therapeutic target in human diseases. Biosci. Rep. 2019; **39**. DOI:10.1042/BSR20190604.

286 Peter B, Matsushita M, Oda K, Raskind W. De novo microdeletion of BCL11A is associated with severe speech sound disorder. *Am J Med Genet A* 2014; **164**. DOI:10.1002/ajmg.a.36599.

287 Tiala I, Suomela S, Huuhtanen J, *et al.* The CCHCR1 (HCR) gene is relevant for skin steroidogenesis and downregulated in cultured psoriatic keratinocytes. *J Mol Med* 2007; **85**. DOI:10.1007/s00109-006-0155-0.

288 Pairo-Castineira E, Clohisey S, Klaric L, *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* 2021; **591**. DOI:10.1038/s41586-020-03065-y.

289 Yao TW, Kim WS, Yu DMT, *et al.* A novel role of dipeptidyl peptidase 9 in epidermal growth factor signaling. *Molecular Cancer Research* 2011; **9**. DOI:10.1158/1541-7786.MCR-10-0272.

290 Sharif-zak M, Abbasi-jorjandi M, Asadikaram G, *et al.* CCR2 and DPP9 expression in the peripheral blood of COVID-19 patients: Influences of the disease severity and gender. *Immunobiology* 2022; **227**. DOI:10.1016/j.imbio.2022.152184.

291 Yuan W, Zhao H, Zhou A, Wang S. Interference of EFNA4 suppresses cell proliferation, invasion and angiogenesis in hepatocellular carcinoma by downregulating PYGO2. *Cancer Biol Ther* 2022; **23**. DOI:10.1080/15384047.2022.2149039.

292 Zhao X, Chen Y, Sun X, *et al.* Oncogenic EFNA4 Amplification Promotes Lung Adenocarcinoma Lymph Node Metastasis. *Cancers (Basel)* 2022; **14**. DOI:10.3390/cancers14174226.

293 El-Agnaf O, Bensmail I, Al-Nesf MAY, *et al.* Uncovering a neurological protein signature for severe COVID-19. *Neurobiol Dis* 2023; **182**. DOI:10.1016/j.nbd.2023.106147.

294 Cerretti DP, Lyman SD, Kozlosky CJ, *et al.* The genes encoding the Eph-related receptor tyrosine kinase ligands LERK-1 (EPLG1, Epl1), LERK-3 (EPLG3, Epl3), and LERK-4 (EPLG4, Epl4) are clustered on

human chromosome 1 and mouse chromosome 3. *Genomics* 1996; **33**. DOI:10.1006/geno.1996.0192.

295 Zhou J, Chehab R, Tkalcevic J, *et al.* Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. *EMBO Journal* 2005; **24**. DOI:10.1038/sj.emboj.7600538.

296 Lee HJ, Ormandy CJ. Elf5, hormones and cell fate. Trends in Endocrinology and Metabolism. 2012; **23**. DOI:10.1016/j.tem.2012.02.006.

297 Perie L, Verma N, Mueller E. The forkhead box transcription factor FoxP4 regulates thermogenic programs in adipocytes. *J Lipid Res* 2021; **62**. DOI:10.1016/J.JLR.2021.100102.

298 Kaur P, Gupta M, Sagar V. FUT2 gene as a genetic susceptible marker of infectious diseases: A Review. *Int J Mol Epidemiol Genet* 2022; **13**.

299 Gómez J, Albaiceta GM, Cuesta-Llavona E, *et al.* The Interferon-induced transmembrane protein 3 gene (IFITM3) rs12252 C variant is associated with COVID-19. *Cytokine* 2021; **137**. DOI:10.1016/j.cyto.2020.155354.

300 Spence JS, He R, Hoffmann HH, *et al.* IFITM3 directly engages and shuttles incoming virus particles to lysosomes. *Nat Chem Biol* 2019; **15**. DOI:10.1038/s41589-018-0213-2.

301 Prelli Bozzo C, Nchioua R, Volcic M, *et al.* IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro. *Nat Commun* 2021; **12**. DOI:10.1038/s41467-021-24817-y.

302 Wellington D, Laurenson-Schafer H, Abdel-Haq A, Dong T. IFITM3: How genetics influence influenza infection demographically. Biomed. J. 2019; **42**. DOI:10.1016/j.bj.2019.01.004.

303 Everitt AR, Clare S, Pertel T, *et al.* IFITM3 restricts the morbidity and mortality associated with influenza. *Nature* 2012; **484**. DOI:10.1038/nature10921.

304 Kim YC, Jeong BH. Strong correlation between the case fatality rate of covid-19 and the rs6598045 single nucleotide polymorphism (Snp) of the interferon-induced transmembrane protein 3 (ifitm3) gene at the population-level. *Genes (Basel)* 2021; **12**. DOI:10.3390/genes12010042.

305 Mulla S, Molla MdMA, Ahmed SMA, Akhtaruzzaman AKM, Saleh AA, Anwar S. Association of interferon gamma inducible protein-10, monocyte chemoattractant protein-1, macrophage inflammatory protein-1 alpha, interleukin-6, and rs12252 single nucleotide polymorphism of interferon-induced transmembrane protein-3 gene with the severity of COVID-19 infection. *Egypt J Intern Med* 2022; **34**. DOI:10.1186/s43162-022-00141-9.

306 Schönfelder K, Breuckmann K, Elsner C, *et al.* The influence of IFITM3 polymorphisms on susceptibility to SARS-CoV-2 infection and severity of COVID-19. *Cytokine* 2021; **142**. DOI:10.1016/j.cyto.2021.155492.

307 Shemesh M, Lochte S, Piehler J, Schreiber G. IFNAR1 and IFNAR2 play distinct roles in initiating type I interferon-induced JAK-STAT signaling and activating STATs. *Sci Signal* 2021; **14**. DOI:10.1126/scisignal.abe4627.

308 Yaugel-Novoa M, Bourlet T, Longet S, Botelho-Nevers E, Paul S. Association of IFNAR1 and IFNAR2 with COVID-19 severity. Lancet Microbe. 2023; **4**. DOI:10.1016/S2666-5247(23)00095-2.

309 Shouval DS, Ouahed J, Biswas A, *et al.* Interleukin 10 receptor signaling: Master regulator of intestinal mucosal homeostasis in mice and humans. In: Advances in Immunology. 2014. DOI:10.1016/B978-0-12-800267-4.00005-5.

310 Ragab D, Salah Eldin H, Taeimah M, Khattab R, Salem R. The COVID-19 Cytokine Storm; What We Know So Far. Front. Immunol. 2020; **11**. DOI:10.3389/fimmu.2020.01446.

311 Kerget B, Kerget F, Koçak AO, *et al.* Are Serum Interleukin 6 and Surfactant Protein D Levels Associated with the Clinical Course of COVID-19? *Lung* 2020; **198**. DOI:10.1007/s00408-020-00393-8.

312 Verma S, Verma S, Khan FH, *et al.* Genetic polymorphisms of IL6 gene –174G > C and –597G > A are associated with the risk of COVID-19 severity. *Int J Immunogenet* 2023; **50**. DOI:10.1111/iji.12605.

313 Yimenicioglu S, Kocaaga A. Koolen-de Vries syndrome: A de novo missense KANSL1 variant. *Clin Neurol Neurosurg* 2022; **222**. DOI:10.1016/j.clineuro.2022.107444.

314 Liu Y, Huang X, Guo L, Luo N. LINC00649 Facilitates the Cellular Process of Bladder Cancer Cells via Signaling Axis miR-16-5p/JARID2. *Urol Int* 2022; **106**. DOI:10.1159/000506239.

315 Si C, Yang L, Cai X. LncRNA LINC00649 aggravates the progression of cervical cancer through sponging miR-216a-3p. *Journal of Obstetrics and Gynaecology Research* 2022; **48**. DOI:10.1111/jog.15405.

316 Huang Q, Li W, Zhou Q, *et al.* Leucine zipper transcription factor-like 1 (LZTFL1), an intraflagellar transporter protein 27 (IFT27) associated protein, is required for normal sperm function and male fertility. *Dev Biol* 2021; **477**. DOI:10.1016/j.ydbio.2021.05.006.

317 Kiss H, Kedra D, Kiss C, *et al.* The LZTFL1 gene is a part of a transcriptional map covering 250 kb within the common eliminated region 1 (C3CER1) in 3p21.3. *Genomics* 2001; **73**. DOI:10.1006/geno.2000.6498.

318 Ridley C, Kouvatsos N, Raynal BD, *et al.* Assembly of the respiratory Mucin MUC5B a new model for a gel-forming Mucin. *Journal of Biological Chemistry* 2014; **289**. DOI:10.1074/jbc.M114.566679.

319 Bonella F, Campo I, Zorzetto M, *et al.* Potential clinical utility of MUC5B und TOLLIP single nucleotide polymorphisms (SNPs) in the management of patients with IPF. *Orphanet J Rare Dis* 2021; **16**. DOI:10.1186/s13023-021-01750-3.

320 van Moorsel CHM, van der Vis JJ, Duckworth A, *et al.* The MUC5B Promoter Polymorphism Associates With Severe COVID-19 in the European Population. *Front Med (Lausanne)* 2021; **8**. DOI:10.3389/fmed.2021.668024.

321 Tian Y, Zhang P, Mou Y, *et al.* Silencing Notch4 promotes tumorigenesis and inhibits metastasis of triple-negative breast cancer via Nanog and Cdc42. *Cell Death Discov* 2023; **9**. DOI:10.1038/s41420-023-01450-w.

322 Harb H, Benamar M, Lai PS, *et al.* Notch4 signaling limits regulatory T-cell-mediated tissue repair and promotes severe lung inflammation in viral infections. *Immunity* 2021; **54**. DOI:10.1016/j.immuni.2021.04.002.

323 Anttila S, Kampman O, Illi A, *et al.* NOTCH4 gene promoter polymorphism is associated with the age of onset in schizophrenia. *Psychiatr Genet* 2003; **13**. DOI:10.1097/01.ypg.0000056681.82896.6b.

324 Schwartz SL, Park EN, Vachon VK, Danzy S, Lowen AC, Conn GL. Human OAS1 activation is highly dependent on both RNA sequence and context of activating RNA motifs. *Nucleic Acids Res* 2020; **48**. DOI:10.1093/nar/gkaa513.

325 Herate C, Ramdani G, Grant NJ, *et al.* Phospholipid scramblase 1 modulates FcR-mediated phagocytosis in differentiated macrophages. *PLoS One* 2016; **11**. DOI:10.1371/journal.pone.0145617.

326 Dal Col J, Lamberti MJ, Nigro A, *et al.* Phospholipid scramblase 1: a protein with multiple functions via multiple molecular interactors. Cell Communication and Signaling. 2022; **20**. DOI:10.1186/s12964-022-00895-3.

327 Babitt JL, Zhang Y, Samad TA, *et al.* Repulsive guidance molecule (RGMa), a DRAGON homologue, is a bone morphogenetic protein co-receptor. *Journal of Biological Chemistry* 2005; **280**. DOI:10.1074/jbc.M503511200.

328 Pizzagalli MD, Bensimon A, Superti-Furga G. A guide to plasma membrane solute carrier proteins. FEBS Journal. 2021; **288**. DOI:10.1111/febs.15531.

329 López-Pacheco C, Soldevila G, Du Pont G, Hernández-Pando R, García-Zepeda EA. CCR9 Is a Key Regulator of Early Phases of Allergic Airway Inflammation. *Mediators Inflamm* 2016; **2016**. DOI:10.1155/2016/3635809.

330 Khan SY, Ali M, Kabir F, *et al.* The role of FYCO1-dependent autophagy in lens fiber cell differentiation. *Autophagy* 2022; **18**. DOI:10.1080/15548627.2022.2025570.

331 Elliott STC, Wetzel KS, Francella N, *et al.* Dualtropic CXCR6/CCR5 Simian Immunodeficiency Virus (SIV) Infection of Sooty Mangabey Primary Lymphocytes: Distinct Coreceptor Use in Natural versus Pathogenic Hosts of SIV. *J Virol* 2015; **89**. DOI:10.1128/jvi.01236-15.

332 Shelton JF, Shastri AJ, Ye C, *et al.* Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat Genet* 2021; **53**. DOI:10.1038/s41588-021-00854-7.

333 Chen Y, Meng H, Meng X, Yan Z, Wang J, Meng F. Correlation between Low THBS3 Expression in Peripheral Blood and Acute Myocardial Infarction. *Frontiers in Bioscience - Landmark* 2022; **27**. DOI:10.31083/j.fbl2710291.

334 Nazli A, Safdar A, Saleem A, *et al.* A mutation in the TMEM65 gene results in mitochondrial myopathy with severe neurological manifestations. *European Journal of Human Genetics* 2017; **25**. DOI:10.1038/ejhg.2017.20.

335 Muromoto R, Oritani K, Matsuda T. Current understanding of the role of tyrosine kinase 2 signaling in immune responses. *World J Biol Chem* 2022; **13**. DOI:10.4331/wjbc.v13.i1.1.

336 Majoros A, Platanitis E, Kernbauer-Hölzl E, Rosebrock F, Müller M, Decker T. Canonical and non-canonical aspects of JAK-STAT signaling:

Lessons from interferons for cytokine responses. Front. Immunol. 2017; **8**. DOI:10.3389/fimmu.2017.00029.

337 Akbari M, Akhavan-Bahabadi M, Shafigh N, *et al.* Expression analysis of IFNAR1 and TYK2 transcripts in COVID-19 patients. *Cytokine* 2022; **153**. DOI:10.1016/j.cyto.2022.155849.

338 Bhowmick R, Gappa-Fahlenkamp H. Cells and Culture Systems Used to Model the Small Airway Epithelium. Lung. 2016; **194**. DOI:10.1007/s00408-016-9875-2.

339 Fulcher ML, Randell SH. Human nasal and tracheo-bronchial respiratory epithelial cell culture. *Methods in Molecular Biology* 2013; **945**. DOI:10.1007/978-1-62703-125-7_8.

340 Chen S, Schoen J. Air-liquid interface cell culture: From airway epithelium to the female reproductive tract. Reproduction in Domestic Animals. 2019; **54**. DOI:10.1111/rda.13481.

341 Kiran KC, Rothenberg ME, Sherrill JD. In vitro model for studying esophageal epithelial differentiation and allergic inflammatory responses identifies keratin involvement in eosinophilic esophagitis. *PLoS One* 2015; **10**. DOI:10.1371/journal.pone.0127755.

342 Kreft ME, Jerman UD, Lasič E, *et al.* The characterization of the human cell line Calu-3 under different culture conditions and its use as an optimized in vitro model to investigate bronchial epithelial function. *European Journal of Pharmaceutical Sciences* 2015. DOI:10.1016/j.ejps.2014.12.017.

343 Han X, Na T, Wu T, Yuan BZ. Human lung epithelial BEAS-2B cells exhibit characteristics of mesenchymal stem cells. *PLoS One* 2020; **15**. DOI:10.1371/journal.pone.0227174.

344 Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 1981; **292**. DOI:10.1038/292154a0.

345 Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* 1981; **78**. DOI:10.1073/pnas.78.12.7634.

346 Thomson JA. Embryonic stem cell lines derived from human blastocysts. *Science (1979)* 1998; **282**. DOI:10.1126/science.282.5391.1145.

347 Cowan CA, Atienza J, Melton DA, Eggan K. Developmental Biology: Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science (1979)* 2005; **309**. DOI:10.1126/science.1116447.

348 Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. *Cell* 2004; **117**. DOI:10.1016/S0092-8674(04)00419-2.

349 Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 2006; **126**. DOI:10.1016/j.cell.2006.07.024.

350 Chronis C, Fiziev P, Papp B, *et al.* Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* 2017; **168**. DOI:10.1016/j.cell.2016.12.016.

351 Zaret KS, Carroll JS. Pioneer transcription factors: Establishing competence for gene expression. Genes Dev. 2011; **25**. DOI:10.1101/gad.176826.111.

352 Vanzan L, Soldati H, Ythier V, *et al.* High throughput screening identifies SOX2 as a super pioneer factor that inhibits DNA methylation maintenance at its binding sites. *Nat Commun* 2021; **12**. DOI:10.1038/s41467-021-23630-x.

353 Takahashi K, Tanabe K, Ohnuki M, *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 2007; **131**. DOI:10.1016/j.cell.2007.11.019.

354 Cerneckis J, Cai H, Shi Y. Induced pluripotent stem cells (iPSCs): molecular mechanisms of induction and applications. *Signal Transduct Target Ther* 2024; **9**: 112.

355 Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: A decade of progress. Nat. Rev. Drug Discov. 2017; **16**. DOI:10.1038/nrd.2016.245.

356 Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. 2009; **10**. DOI:10.1038/nrg2484.

357 Kukurba KR, Montgomery SB. RNA sequencing and analysis. *Cold Spring Harb Protoc* 2015; **2015**. DOI:10.1101/pdb.top084970.

358 Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: A matter of depth. *Genome Res* 2011; **21**. DOI:10.1101/gr.124321.111.

359 Rai MF, Tycksen ED, Sandell LJ, Brophy RH. Advantages of RNA-seq compared to RNA microarrays for transcriptome profiling of anterior cruciate ligament tears. *Journal of Orthopaedic Research* 2018; **36**. DOI:10.1002/jor.23661.

360 Wiltgen M, Tilz GP. Molecular diagnosis and prognosis with DNA micro arrays. *Hematology* 2011; **16**. DOI:10.1179/102453311X12953015767257.

361 Bumgarner R. Overview of dna microarrays: Types, applications, and their future. *Curr Protoc Mol Biol* 2013. DOI:10.1002/0471142727.mb2201s101.

362 Wang Q, Wang Z, Zhang J, *et al.* Dual RNA-Seq of H5N1 Avian Influenza Virus and Host Cell Transcriptomes Reveals Novel Insights Into Host-Pathogen Cross Talk. *Front Microbiol* 2022; **13**. DOI:10.3389/fmicb.2022.828277.

363 Liang W, Wang P, Wang X, Xu D, Chunfu Zheng. Application of Single-Cell RNA Sequencing in Investigating Virus–Host Interactions. *Methods in Molecular Biology*.

364 Mazen NA, Luc J, Lobb B, Hirota JA, Banerjee A, Doxey AC. Metatranscriptomic RNA-Seq Data Analysis of Virus-Infected Host Cells. In: Methods in Molecular Biology. .

365 Rodríguez A, Duyvejonck H, van Belleghem JD, *et al.* Comparison of procedures for RNA-extraction from peripheral blood mononuclear cells. *PLoS One* 2020; **15**. DOI:10.1371/journal.pone.0229423.

366 Philippeos C, Hughes RD, Dhawan A, Mitry RR. Introduction to cell culture. *Methods in Molecular Biology* 2012. DOI:10.1007/978-1-61779-367-7_1.

367 Chua PF, Lim WK. The strategic uses of collagen in adherent cell cultures. *Cell Biol Int* 2023; **47**. DOI:10.1002/cbin.11966.

368 Tan E, Chin CSH, Lim ZFS, Ng SK. HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. Front. Bioeng. Biotechnol. 2021; **9**. DOI:10.3389/fbioe.2021.796991.

369 Hallak LK, Spillmann D, Collins PL, Peeples ME. Glycosaminoglycan Sulfation Requirements for Respiratory Syncytial Virus Infection. *J Virol* 2000; **74**. DOI:10.1128/jvi.74.22.10508-10513.2000.

370 Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat. Rev. Genet. 2019; **20**. DOI:10.1038/s41576-019-0127-1.

371 Murthy A, Li Y, Peng I, *et al.* A Crohn's disease variant in Atg16l1 enhances its degradation by caspase 3. *Nature* 2014; **506**. DOI:10.1038/nature13044.

372 Doyle WJ, Casselbrant ML, Li-Korotky HS, *et al.* The interleukin 6-174 C/C genotype predicts greater rhinovirus illness. *Journal of Infectious Diseases* 2010; **201**. DOI:10.1086/649559.

373 Alvarez AE, Marson FAL, Bertuzzo CS, *et al.* Association between single nucleotide polymorphisms in TLR4, TLR2, TLR9, VDR, NOS2 and CCL5 genes with acute viral bronchiolitis. *Gene* 2018; **645**. DOI:10.1016/j.gene.2017.12.022.

374 Salas A, Pardo-Seco J, Cebey-López M, *et al.* Whole Exome Sequencing reveals new candidate genes in host genomic susceptibility to Respiratory Syncytial Virus Disease. *Sci Rep* 2017; **7**. DOI:10.1038/s41598-017-15752-4.

375 Mahajan MC, McLellan AS. Whole-exome sequencing (WES) for illumina short read sequencers using solution-based capture. In: Methods in Molecular Biology. 2020. DOI:10.1007/978-1-4939-9882-1_5.

376 D E, F D, L B, *et al.* Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *New England Journal of Medicine* 2020; **383**.

377 Niemi MEK, Karjalainen J, Liao RG, *et al.* Mapping the human genetic architecture of COVID-19. *Nature* 2021; **600**: 472–7.

378 Pairo-Castineira E, Clohisey S, Klaric L, *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* 2021; **591**. DOI:10.1038/s41586-020-03065-y.

379 Hubacek JA, Dusek L, Majek O, *et al.* ACE I/D polymorphism in Czech first-wave SARS-CoV-2-positive survivors. *Clinica Chimica Acta* 2021; **519**. DOI:10.1016/j.cca.2021.04.024.

380 Verma S, Abbas M, Verma S, *et al.* Impact of I/D polymorphism of angiotensin-converting enzyme 1 (ACE1) gene on the severity of COVID-19 patients. *Infection, Genetics and Evolution* 2021; **91**. DOI:10.1016/j.meegid.2021.104801.

381 Soltani Rezaiezadeh J, Lord JS, Yekaninejad MS, Izadi P. The association of ACE I/D polymorphism with the severity of COVID-19 in Iranian patients: A case-control study. *Human Gene* 2022; **34**. DOI:10.1016/j.humgen.2022.201099.

382 Möhlendick B, Schönfelder K, Breuckmann K, *et al.* ACE2 polymorphism and susceptibility for SARS-CoV-2 infection and severity of COVID-19. *Pharmacogenet Genomics* 2021. DOI:10.1097/FPC.0000000000000436.

383 Schönfelder K, Breuckmann K, Elsner C, *et al.* Transmembrane serine protease 2 Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Type 2 Infection: A German Case-Control Study. *Front Genet* 2021; **12**. DOI:10.3389/fgene.2021.667231.

384 Mulla S, Molla MdMA, Ahmed SMA, Akhtaruzzaman AKM, Saleh AA, Anwar S. Association of interferon gamma inducible protein-10, monocyte chemoattractant protein-1, macrophage inflammatory protein-1 alpha, interleukin-6, and rs12252 single nucleotide polymorphism of interferon-induced transmembrane protein-3 gene with the severity of COVID-19 infection. *Egypt J Intern Med* 2022; **34**. DOI:10.1186/s43162-022-00141-9.

385 Verma S, Verma S, Khan FH, *et al.* Genetic polymorphisms of IL6 gene –174G > C and –597G > A are associated with the risk of COVID-19 severity. *Int J Immunogenet* 2023; **50**. DOI:10.1111/iji.12605.

386 Heidari Nia M, Rokni M, Mirinejad S, *et al.* Association of polymorphisms in tumor necrosis factors with SARS-CoV-2 infection and mortality rate: A case-control study and in silico analyses. *J Med Virol* 2022; **94**. DOI:10.1002/jmv.27477.

387 Saleh A, Sultan A, Elashry M a., *et al.* Association of TNF-α G-308 a Promoter Polymorphism with the Course and Outcome of COVID-19 Patients. *Immunol Invest* 2020. DOI:10.1080/08820139.2020.1851709.

388 Espinosa-Salinas I, Colmenarejo G, Fernández-Díaz CM, *et al.* Potential protective effect against SARS-CoV-2 infection by APOE rs7412 polymorphism. *Sci Rep* 2022; **12**. DOI:10.1038/s41598-022-10923-4.

389 van Moorsel CHM, van der Vis JJ, Duckworth A, *et al.* The MUC5B Promoter Polymorphism Associates With Severe COVID-19 in the European Population. *Front Med (Lausanne)* 2021; **8**. DOI:10.3389/fmed.2021.668024.

390 Nelson MR, Tipney H, Painter JL, *et al.* The support of human genetic evidence for approved drug indications. *Nat Genet* 2015; **47**. DOI:10.1038/ng.3314.

391 Polack FP, Thomas SJ, Kitchin N, *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* 2020; **383**. DOI:10.1056/nejmoa2034577.

392 Baden LR, El Sahly HM, Essink B, *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine* 2021; **384**. DOI:10.1056/nejmoa2035389.

393 Heath PT, Galiza EP, Baxter DN, *et al.* Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine. *New England Journal of Medicine* 2021; **385**. DOI:10.1056/nejmoa2107659.

394 Li G, Hilgenfeld R, Whitley R, De Clercq E. Therapeutic strategies for COVID-19: progress and lessons learned. Nat. Rev. Drug Discov. 2023; **22**. DOI:10.1038/s41573-023-00672-y.

395 Li L, Zhang X, Zhao H. EQTL. *Methods in Molecular Biology* 2012; **871**. DOI:10.1007/978-1-61779-785-9_14.

396 Aguet F, Barbeira AN, Bonazzola R, *et al.* The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science (1979)* 2020; **369**. DOI:10.1126/SCIENCE.AAZ1776.

397 Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**. DOI:10.1093/nar/gkr917.

398 Ghoussaini M, Mountjoy E, Carmona M, *et al.* Open Targets Genetics: Systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res* 2021; **49**. DOI:10.1093/nar/gkaa840.

399 Shaw DE, Sousa AR, Fowler SJ, *et al.* Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma

cohort. In: European Respiratory Journal. 2015. DOI:10.1183/13993003.00779-2015.

400 Slatkin M. Linkage disequilibrium - Understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. 2008; **9**. DOI:10.1038/nrg2361.

401 Machiela MJ, Chanock SJ. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015; **31**. DOI:10.1093/bioinformatics/btv402.

402 Buniello A, Macarthur JAL, Cerezo M, *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019; **47**. DOI:10.1093/nar/gky1120.

403 Sollis E, Mosaku A, Abid A, *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* 2023; **51**. DOI:10.1093/nar/gkac1010.

404 Bastarache L, Denny JC, Roden DM. Phenome-Wide Association Studies. *JAMA* 2022; **327**. DOI:10.1001/jama.2021.20356.

405 Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet* 2018; **50**: 1593–9.

406 Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. *Protein Science* 2018; **27**. DOI:10.1002/pro.3307.

407 Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 2014; **30**. DOI:10.1093/bioinformatics/btt703.

408 Freshour SL, Kiwala S, Cotto KC, *et al.* Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res* 2021; **49**. DOI:10.1093/nar/gkaa1084.

409 Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov Results Database — Update and Key Issues. *New England Journal of Medicine* 2011; **364**. DOI:10.1056/nejmsa1012065.

410 Portelli MA, Rakkar K, Hu S, Guo Y, Adcock IM, Sayers I. Translational Analysis of Moderate to Severe Asthma GWAS Signals Into Candidate Causal Genes and Their Functional, Tissue-Dependent and Disease-Related Associations. *Frontiers in Allergy* 2021; **2**. DOI:10.3389/falgy.2021.738741.

411 Gong B-S, Zhang Q-P, Zhang G-M, *et al.* Single-nucleotide polymorphism-gene intermixed networking reveals co-linkers connected to multiple gene expression phenotypes. *BMC Proc* 2007; **1**. DOI:10.1186/1753-6561-1-s1-s45.

412 Spitaels J, Van Hoecke L, Roose K, Kochs G, Saelens X. Mx1 in Hematopoietic Cells Protects against Thogoto Virus Infection. *J Virol* 2019; **93**. DOI:10.1128/jvi.00193-19.

413 Tian Y, Jain S, Kelemen SE, Autieri M V. AIF-1 expression regulates endothelial cell activation, signal transduction, and vasculogenesis. *Am J Physiol Cell Physiol* 2009; **296**. DOI:10.1152/ajpcell.00325.2008.

414 Tiala I, Suomela S, Huuhtanen J, *et al.* The CCHCR1 (HCR) gene is relevant for skin steroidogenesis and downregulated in cultured psoriatic keratinocytes. *J Mol Med* 2007; **85**. DOI:10.1007/s00109-006-0155-0.

415 Awasthi S, Chakrapani B, Mahesh A, Chavali PL, Chavali S, Dhayalan A. DDX39B promotes translation through regulation of pre-ribosomal RNA levels. *RNA Biol* 2018; **15**. DOI:10.1080/15476286.2018.1517011.

416 Poujade FA, Mannion A, Brittain N, *et al.* WSB-1 regulates the metastatic potential of hormone receptor negative breast cancer. *Br J Cancer* 2018; **118**. DOI:10.1038/s41416-018-0056-3.

417 Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically

Proven Protease Inhibitor. *Cell* 2020; **181**. DOI:10.1016/j.cell.2020.02.052.

418 de Melo Vilar K, Pereira MC, Dantas AT, *et al.* Galectin-9 gene (LGALS9) polymorphisms are associated with rheumatoid arthritis in Brazilian patients. *PLoS One* 2019; **14**. DOI:10.1371/journal.pone.0223191.

419 Bruni F, Di Meo I, Bellacchio E, *et al.* Clinical, biochemical, and genetic features associated with VARS2-related mitochondrial disease. *Hum Mutat* 2018; **39**. DOI:10.1002/humu.23398.

420 Holgate ST, Davies DE, Powell RM, Holloway JW. ADAM33: A newly identified protease involved in airway remodelling. *Pulm Pharmacol Ther* 2006; **19**. DOI:10.1016/j.pupt.2005.02.008.

421 Hadchouel A, Decobert F, Franco-Montoya ML, *et al.* Matrix metalloproteinase gene polymorphisms and bronchopulmonary dysplasia: Identification of MMP16 as a new player in lung development. *PLoS One* 2008; **3**. DOI:10.1371/journal.pone.0003188.

422 Pardo A, Selman M. MMP-1: The elder of the family. International Journal of Biochemistry and Cell Biology. 2005; **37**. DOI:10.1016/j.biocel.2004.06.017.

423 Nahm M, Lim SM, Kim YE, *et al.* ANXA11 mutations in ALS cause dysregulation of calcium homeostasis and stress granule dynamics. *Sci Transl Med* 2020; **12**. DOI:10.1126/scitranslmed.aax3993.

424 Kasu YAT, Arva A, Johnson J, *et al.* BAG6 prevents the aggregation of neurodegeneration-associated fragments of TDP43. *iScience* 2022; **25**. DOI:10.1016/j.isci.2022.104273.

425 Abdulhag UN, Soiferman D, Schueler-Furman O, *et al.* Mitochondrial complex IV deficiency, caused by mutated COX6B1, is associated with encephalomyopathy, hydrocephalus and cardiomyopathy. *European Journal of Human Genetics* 2015; **23**. DOI:10.1038/ejhg.2014.85.

426 Wu L, Lu P, Guo X, *et al.* β-carotene oxygenase 2 deficiency-triggered mitochondrial oxidative stress promotes low-grade inflammation and metabolic dysfunction. *Free Radic Biol Med* 2021; **164**. DOI:10.1016/j.freeradbiomed.2021.01.003.

427 Roberts RM, Ezashi T, Temple J, Owen JR, Soncin F, Parast MM. The role of BMP4 signaling in trophoblast emergence from pluripotency. Cellular and Molecular Life Sciences. 2022; **79**. DOI:10.1007/s00018-022-04478-w.

428 Stasia MJ. CYBA encoding p22phox, the cytochrome b558 alpha polypeptide: Gene structure, expression, role and physiopathology. Gene. 2016; **586**. DOI:10.1016/j.gene.2016.03.050.

429 Ardehali MB, Damle M, Perea-Resa C, Blower MD, Kingston RE. Elongin A associates with actively transcribed genes and modulates enhancer RNA levels with limited impact on transcription elongation rate in vivo. *Journal of Biological Chemistry* 2021; **296**. DOI:10.1074/JBC.RA120.015877.

430 Kim KH, Hong EP, Shin JW, *et al.* Genetic and Functional Analyses Point to FAN1 as the Source of Multiple Huntington Disease Modifier Effects. *Am J Hum Genet* 2020; **107**. DOI:10.1016/j.ajhg.2020.05.012.

431 Völkening L, Vatselia A, Asgedom G, *et al.* RAD50 regulates mitotic progression independent of DNA repair functions. *FASEB Journal* 2020; **34**. DOI:10.1096/fj.201902318R.

432 Heath RJ, Goel G, Baxt LA, *et al.* RNF166 Determines Recruitment of Adaptor Proteins during Antibacterial Autophagy. *Cell Rep* 2016; **17**. DOI:10.1016/j.celrep.2016.11.005.

433 Hwang IY, Oh CK, Choi YK, Yun N, Oh YJ. RNF166 plays a dual role for Lys63-linked ubiquitination and sumoylation of its target proteins. *J Neural Transm* 2022; **129**. DOI:10.1007/s00702-021-02442-9.

434 Keele GR, Prokop JW, He H, *et al.* Sept8/SEPTIN8 involvement in cellular structure and kidney damage is identified by genetic mapping

and a novel human tubule hypoxic model. *Sci Rep* 2021; **11**. DOI:10.1038/s41598-021-81550-8.

435 Fukuhara N, Ebert J, Unterholzner L, Lindner D, Izaurralde E, Conti E. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol Cell* 2005; **17**. DOI:10.1016/j.molcel.2005.01.010.

436 Leygue E. Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. Nucl. Recept. Signal. 2007; **5**. DOI:10.1621/nrs.05006.

437 Sellegounder D, Zafari P, Rajabinejad M, Taghadosi M, Kapahi P. Advanced glycation end products (AGEs) and its receptor, RAGE, modulate age-dependent COVID-19 morbidity and mortality. A review and hypothesis. Int. Immunopharmacol. 2021; **98**. DOI:10.1016/j.intimp.2021.107806.

438 Yamazaki Y, Zhao N, Caulfield TR, Liu CC, Bu G. Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. Nat. Rev. Neurol. 2019; **15**. DOI:10.1038/s41582-019-0228-7.

439 Novikova G, Kapoor M, Tcw J, *et al.* Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. *Nat Commun* 2021; **12**. DOI:10.1038/s41467-021-21823-y.

440 Yin J, Xie X, Ye Y, Wang L, Che F. BCL11A: A potential diagnostic biomarker and therapeutic target in human diseases. Biosci. Rep. 2019; **39**. DOI:10.1042/BSR20190604.

441 Smith EC, Luc S, Croney DM, *et al.* Strict in vivo specificity of the Bcl11a erythroid enhancer. *Blood* 2016; **128**. DOI:10.1182/blood-2016-08-736249.

442 Bhalla K, Luo Y, Buchan T, *et al.* Alterations in CDH15 and KIRREL3 in Patients with Mild to Severe Intellectual Disability. *Am J Hum Genet* 2008; **83**. DOI:10.1016/j.ajhg.2008.10.020.

443 Pei YP, Wang YY, Liu D, *et al.* ICAM5 as a novel target for treating cognitive impairment in fragile X syndrome. *Journal of Neuroscience* 2020; **40**. DOI:10.1523/JNEUROSCI.2626-18.2019.

444 Zhang Y, Willis-Owen SAG, Spiegel S, Lloyd CM, Moffatt MF, Cookson WOCM. The ORMDL3 asthma gene regulates ICAM1 and has multiple effects on cellular inflammation. *Am J Respir Crit Care Med* 2019; **199**. DOI:10.1164/rccm.201803-0438OC.

445 Huang S, Li Z, Wu Z, *et al.* DDAH2 suppresses RLR-MAVS mediated innate antiviral immunity by stimulating nitric oxide activated, Drp1-induced mitochondrial fission. *Sci Signal* 2021; **14**. DOI:10.1126/scisignal.abc7931.

446 Nachiyappan A, Gupta N, Taneja R. EHMT1/EHMT2 in EMT, cancer stemness and drug resistance: emerging evidence and mechanisms. FEBS Journal. 2022; **289**. DOI:10.1111/febs.16334.

447 Koo Y, Barry DM, Xu K, *et al.* Rasip1 is essential to blood vessel stability and angiogenic blood vessel growth. *Angiogenesis* 2016; **19**. DOI:10.1007/s10456-016-9498-5.

448 Hüttelmaier S, Illenberger S, Grosheva I, Rüdiger M, Singer RH, Jockusch BM. Raver1, a dual compartment protein, is a ligand for PTB/hnRNPI and microfilament attachment proteins. *Journal of Cell Biology* 2001; **155**. DOI:10.1083/jcb.200105044.

449 Mondal P, Gadad SS, Adhikari S, *et al.* TCF19 and p53 regulate transcription of TIGAR and SCO2 in HCC for mitochondrial energy metabolism and stress adaptation. *FASEB Journal* 2021; **35**. DOI:10.1096/fj.202002486RR.

450 Pénisson-Besnier I, Allamand V, Beurrier P, *et al.* Compound heterozygous mutations of the TNXB gene cause primary myopathy. *Neuromuscular Disorders* 2013; **23**. DOI:10.1016/j.nmd.2013.04.009.

451 Krivega M, Essahib W, Van de Velde H. WNT3 and membrane-associated β-catenin regulate trophectoderm lineage differentiation in

human blastocysts. *Mol Hum Reprod* 2015; **21**. DOI:10.1093/molehr/gav036.

452 Jiang W, Shi J, Zhao J, *et al.* ZFP57 dictates allelic expression switch of target imprinted genes. *Proc Natl Acad Sci U S A* 2021; **118**. DOI:10.1073/pnas.2005377118.

453 Le Rolle M, Massa F, Siggers P, *et al.* Arrest of WNT/β-catenin signaling enables the transition from pluripotent to differentiated germ cells in mouse ovaries. *Proc Natl Acad Sci U S A* 2021; **118**. DOI:10.1073/pnas.2023376118.

454 Yi E, Zhang J, Zheng M, *et al.* Long noncoding RNA IL6-AS1 is highly expressed in chronic obstructive pulmonary disease and is associated with interleukin 6 by targeting miR-149-5p and early B-cell factor 1. *Clin Transl Med* 2021; **11**. DOI:10.1002/ctm2.479.

455 Huang N, Zhang J, Kuang S, *et al.* Role of NCF2 as a potential prognostic factor and immune infiltration indicator in hepatocellular carcinoma. *Cancer Med* 2023; **12**. DOI:10.1002/cam4.5597.

456 Zhu Z, Tang W, Ray A, *et al.* Rhinovirus stimulation of Interleukin-6 in vivo and in vitro: Evidence for nuclear factor κB-dependent transcriptional activation. *Journal of Clinical Investigation* 1996; **97**. DOI:10.1172/JCI118431.

457 Santa Cruz A, Mendes-Frias A, Oliveira AI, *et al.* Interleukin-6 Is a Biomarker for the Development of Fatal Severe Acute Respiratory Syndrome Coronavirus 2 Pneumonia. *Front Immunol* 2021; **12**. DOI:10.3389/fimmu.2021.613422.

458 Barry DM, Koo Y, Norden PR, *et al.* Rasip1-Mediated Rho GTPase Signaling Regulates Blood Vessel Tubulogenesis via Nonmuscle Myosin II. *Circ Res* 2016; **119**. DOI:10.1161/CIRCRESAHA.116.309094.

459 Muehling LM, Heymann PW, Wright PW, *et al.* Human TH1 and TH2 cells targeting rhinovirus and allergen coordinately promote allergic

asthma. *Journal of Allergy and Clinical Immunology* 2020; **146**. DOI:10.1016/j.jaci.2020.03.037.

460 Makino A, Shibata T, Nagayasu M, *et al.* RSV infection-elicited high MMP-12–producing macrophages exacerbate allergic airway inflammation with neutrophil infiltration. *iScience* 2021; **24**. DOI:10.1016/j.isci.2021.103201.

461 Ravi LI, Li L, Sutejo R, *et al.* A systems-based approach to analyse the host response in murine lung macrophages challenged with respiratory syncytial virus. *BMC Genomics* 2013; **14**. DOI:10.1186/1471-2164-14-190.

462 Yang L, Xie X, Tu Z, Fu J, Xu D, Zhou Y. The signal pathways and treatment of cytokine storm in COVID-19. Signal Transduct. Target. Ther. 2021; **6**. DOI:10.1038/s41392-021-00679-0.

463 Zhao Y, Qin L, Zhang P, *et al.* Longitudinal COVID-19 profiling associates IL-1RA and IL-10 with disease severity and RANTES with mild disease. *JCI Insight* 2020; **5**. DOI:10.1172/jci.insight.139834.

464 Teng S, Yu M, Qi Z, Zhao S, Wang H. Expression analysis of the IL-23/Th17 pathway in children with rhinovirus infection. *Ital J Pediatr* 2025.

465 Hidalgo A, Libby P, Soehnlein O, Aramburu IV, Papayannopoulos V, Silvestre-Roig C. Neutrophil extracellular traps: from physiology to pathology. Cardiovasc. Res. 2022; **118**. DOI:10.1093/cvr/cvab329.

466 Cortjens B, De Boer OJ, De Jong R, *et al.* Neutrophil extracellular traps cause airway obstruction during respiratory syncytial virus disease. *Journal of Pathology* 2016; **238**. DOI:10.1002/path.4660.

467 Allenby G, Bocquel MT, Saunders M, *et al.* Retinoic acid receptors and retinoid X receptors: Interactions with endogenous retinoic acids. *Proc Natl Acad Sci U S A* 1993; **90**. DOI:10.1073/pnas.90.1.30.

468 Halvorsen B, Viermyr H-K, Ueland T, *et al.* IL-22 is increased in hospitalized patients with COVID-19 and associates with cardiac involvement. *Journal of Infection* 2024.

469 Hanna S, Tibby SM, Durward A, Murdoch IA. Incidence of hyponatraemia and hyponatraemic seizures in severe respiratory syncytial virus bronchiolitis. Acta Paediatrica, International Journal of Paediatrics. 2003; **92**. DOI:10.1111/j.1651-2227.2003.tb00573.x.

470 Puig-Domingo M, Marazuela M, Yildiz BO, Giustina A. COVID-19 and endocrine and metabolic diseases. An updated statement from the European Society of Endocrinology. *Endocrine* 2021; **72**. DOI:10.1007/s12020-021-02734-w.

471 Li J, Wang X, Chen J, Zuo X, Zhang H, Deng A. COVID-19 infection may cause ketosis and ketoacidosis. *Diabetes Obes Metab* 2020; **22**. DOI:10.1111/dom.14057.

472 Rubino F, Amiel SA, Zimmet P, *et al.* New-Onset Diabetes in Covid-19. *New England Journal of Medicine* 2020; **383**. DOI:10.1056/nejmc2018688.

473 Zhang J, Garrett S, Sun J. Gastrointestinal symptoms, pathophysiology, and treatment in COVID-19. Genes Dis. 2021; **8**. DOI:10.1016/j.gendis.2020.08.013.

474 Hazama K, Shiihara T, Tsukagoshi H, Matsushige T, Dowa Y, Watanabe M. Rhinovirus-associated acute encephalitis/encephalopathy and cerebellitis. *Brain Dev* 2019; **41**. DOI:10.1016/j.braindev.2019.02.014.

475 Saravanos GL, King CL, Deng L, *et al.* Respiratory Syncytial Virus–Associated Neurologic Complications in Children: A Systematic Review and Aggregated Case Series. *Journal of Pediatrics* 2021; **239**. DOI:10.1016/j.jpeds.2021.06.045.

476 Jiang Y, Mao X, Song R, Mu K, Yang Y, Zhang J an. Psoriasis Susceptibility 1 Candidate 1 (PSORS1C1) Polymorphism is Associated

with Autoimmune Thyroid Disease in a Chinese Han Population. *Immunol Invest* 2022; **51**. DOI:10.1080/08820139.2021.1923734.

477 Maury A, Lyoubi A, Peiffer-Smadja N, de Broucker T, Meppiel E. Neurological manifestations associated with SARS-CoV-2 and other coronaviruses: A narrative review for clinicians. Rev. Neurol. (Paris). 2021; **177**. DOI:10.1016/j.neurol.2020.10.001.

478 Lim SH, Ju HJ, Han JH, *et al.* Autoimmune and Autoinflammatory Connective Tissue Disorders Following COVID-19. *JAMA Netw Open* 2023; **6**. DOI:10.1001/jamanetworkopen.2023.36120.

479 Groneberg DA, Folkerts G, Peiser C, Chung KF, Fischer A. Neuropeptide Y (NPY). *Pulm Pharmacol Ther* 2004; **17**. DOI:10.1016/j.pupt.2004.04.003.

480 Decressac M, Barker RA. Neuropeptide Y and its role in CNS disease and repair. Exp. Neurol. 2012; **238**. DOI:10.1016/j.expneurol.2012.09.004.

481 Monick MM, Yarovinsky TO, Powers LS, *et al.* Respiratory Syncytial Virus Up-regulates TLR4 and Sensitizes Airway Epithelial Cells to Endotoxin. *Journal of Biological Chemistry* 2003; **278**. DOI:10.1074/jbc.M308093200.

482 Rallabhandi P, Phillips RL, Boukhvalova MS, *et al.* Respiratory syncytial virus fusion protein-induced toll-like receptor 4 (TLR4) signaling is inhibited by the TLR4 antagonists rhodobacter sphaeroides lipopolysaccharide and eritoran (E5564) and requires direct interaction with MD-2. *mBio* 2012; **3**. DOI:10.1128/mBio.00218-12.

483 Garza Flores A, Nordgren I, Pettersson M, *et al.* Case report: Extending the spectrum of clinical and molecular findings in FOXC1 haploinsufficiency syndrome. *Front Genet* 2023; **14**. DOI:10.3389/fgene.2023.1174046.

484 Manning BD, Toker A. AKT/PKB Signaling: Navigating the Network. Cell. 2017; **169**. DOI:10.1016/j.cell.2017.04.001.

485 Bentley JK, Newcomb DC, Goldsmith AM, Jia Y, Sajjan US, Hershenson MB. Rhinovirus Activates Interleukin-8 Expression via a Src/p110β Phosphatidylinositol 3-Kinase/Akt Pathway in Human Airway Epithelial Cells. *J Virol* 2007; **81**. DOI:10.1128/jvi.02309-06.

486 Massagué J. TGFβ signalling in context. Nat. Rev. Mol. Cell Biol. 2012; **13**. DOI:10.1038/nrm3434.

487 Gibbs JD, Ornoff DM, Igo HA, Zeng JY, Imani F. Cell Cycle Arrest by Transforming Growth Factor β1 Enhances Replication of Respiratory Syncytial Virus in Lung Epithelial Cells. *J Virol* 2009; **83**. DOI:10.1128/jvi.00806-09.

488 Codo AC, Davanzo GG, Monteiro L de B, *et al.* Elevated Glucose Levels Favor SARS-CoV-2 Infection and Monocyte Response through a HIF-1α/Glycolysis-Dependent Axis. *Cell Metab* 2020; **32**. DOI:10.1016/j.cmet.2020.07.007.

489 Zou M, Hu X, Song W, *et al.* Plasma LTBP2 as a potential biomarker in differential diagnosis of connective tissue disease-associated interstitial lung disease and idiopathic pulmonary fibrosis: a pilot study. *Clin Exp Med* 2023; **23**. DOI:10.1007/s10238-023-01214-x.

490 Lee JJ, Ishihara K, Notomi S, *et al.* Lysosome-associated membrane protein-2 deficiency increases the risk of reactive oxygen species-induced ferroptosis in retinal pigment epithelial cells. *Biochem Biophys Res Commun* 2020; **521**. DOI:10.1016/j.bbrc.2019.10.138.

491 Valkov N, Das S. Y RNAs: Biogenesis, Function and Implications for the Cardiovascular System. In: Advances in Experimental Medicine and Biology. 2020. DOI:10.1007/978-981-15-1671-9_20.

492 Estravís M, García-Sánchez A, Martin MJ, *et al.* RNY3 modulates cell proliferation and IL13 mRNA levels in a T lymphocyte model: a possible new epigenetic mechanism of IL-13 regulation. *J Physiol Biochem* 2023; **79**. DOI:10.1007/s13105-022-00920-6.

493 Floris M, Olla S, Schlessinger D, Cucca F. Genetic-Driven Druggable Target Identification and Validation. Trends in Genetics. 2018; **34**. DOI:10.1016/j.tig.2018.04.004.

494 Mease PJ, Gottlieb AB, Berman A, *et al.* The Efficacy and Safety of Clazakizumab, an Anti–Interleukin-6 Monoclonal Antibody, in a Phase IIb Study of Adults With Active Psoriatic Arthritis. *Arthritis and Rheumatology* 2016; **68**. DOI:10.1002/art.39700.

495 Lonze BE, Spiegler P, Wesson RN, *et al.* A Randomized Double-Blinded Placebo Controlled Trial of Clazakizumab for the Treatment of COVID-19 Pneumonia With Hyperinflammation. *Crit Care Med* 2022; **50**. DOI:10.1097/CCM.0000000000005591.

496 Glanville N, Johnston SL. Challenges in developing a cross-serotype rhinovirus vaccine. Curr. Opin. Virol. 2015; **11**. DOI:10.1016/j.coviro.2015.03.004.

497 Li J, Varghese RS, Ressom HW. RNA-Seq Data Analysis. *Methods Mol Biol* 2024; **2822**: 263–90.

498 Liu R, Holik AZ, Su S, *et al.* Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Res* 2015; **43**. DOI:10.1093/nar/gkv412.

499 Nicolas de Lamballerie C, Pizzorno A, Dubois J, *et al.* Characterization of cellular transcriptomic signatures induced by different respiratory viruses in human reconstituted airway epithelia. *Sci Rep* 2019; **9**. DOI:10.1038/s41598-019-48013-7.

500 Parkhomchuk D, Borodina T, Amstislavskiy V, *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res* 2009; **37**. DOI:10.1093/nar/gkp596.

501 Yick CY, Zwinderman AH, Kunst PW, *et al.* Transcriptome sequencing (RNA-Seq) of human endobronchial biopsies: Asthma versus controls. *European Respiratory Journal* 2013; **42**. DOI:10.1183/09031936.00115412.

502 Morrow JD, Chase RP, Parker MM, *et al.* RNA-sequencing across three matched tissues reveals shared and tissue-specific gene expression and pathway signatures of COPD. *Respir Res* 2019; **20**. DOI:10.1186/s12931-019-1032-z.

503 Chappin K, Besteman SB, Hennus MP, *et al.* Airway and Blood Monocyte Transcriptomic Profiling Reveals an Antiviral Phenotype in Infants With Severe Respiratory Syncytial Virus Infection. *Journal of Infectious Diseases* 2024; **229**. DOI:10.1093/infdis/jiad487.

504 Wang L, Chu CY, McCall MN, *et al.* Airway gene-expression classifiers for respiratory syncytial virus (RSV) disease severity in infants. *BMC Med Genomics* 2021; **14**. DOI:10.1186/s12920-021-00913-2.

505 Rai KR, Shrestha P, Yang B, *et al.* Acute Infection of Viral Pathogens and Their Innate Immune Escape. Front. Microbiol. 2021; **12**. DOI:10.3389/fmicb.2021.672026.

506 Gwaltney JM. Clinical significance and pathogenesis of viral respiratory infections. In: American Journal of Medicine. 2002. DOI:10.1016/s0002-9343(01)01059-2.

507 Ong HH, Andiappan AK, Duan K, *et al.* Transcriptomics of rhinovirus persistence reveals sustained expression of RIG-I and interferon-stimulated genes in nasal epithelial cells in vitro. *Allergy: European Journal of Allergy and Clinical Immunology* 2022; **77**. DOI:10.1111/all.15280.

508 Cai G, Li H, Lu Y, *et al.* Accuracy of RNA-Seq and its dependence on sequencing depth. *BMC Bioinformatics* 2012; **13 Suppl 13**. DOI:10.1186/1471-2105-13-S13-S5.

509 Li S, Liu J, Zhao M, Su Y, Cong B, Wang Z. RNA quality score evaluation: A preliminary study of RNA integrity number (RIN) and RNA integrity and quality number (RNA IQ). *Forensic Sci Int* 2024; **357**. DOI:10.1016/j.forsciint.2024.111976.

510 Sheng Q, Vickers K, Zhao S, *et al.* Multi-perspective quality control of Illumina RNA sequencing data analysis. *Brief Funct Genomics* 2017; **16**. DOI:10.1093/bfgp/elw035.

511 Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. An extensive evaluation of read trimming effects on illumina NGS data analysis. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0085024.

512 Rajasekar P, Hall RJ, Binaya KC, *et al.* Nepalese indoor cookstove smoke extracts alter human airway epithelial gene expression, DNA methylation and hydroxymethylation. *Environmental Pollution* 2023; **337**. DOI:10.1016/j.envpol.2023.122561.

513 Sewe SO, Silva G, Sicat P, Seal SE, Visendi P. Trimming and Validation of Illumina Short Reads Using Trimmomatic, Trinity Assembly, and Assessment of RNA-Seq Data. In: Methods in Molecular Biology. 2022. DOI:10.1007/978-1-0716-2067-0_11.

514 Dobin A, Davis CA, Schlesinger F, *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**. DOI:10.1093/bioinformatics/bts635.

515 Oliveros JC. Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams. - References - Scientific Research Publishing. Csic. 2007.

516 Smyth GK, Ritchie ME, Law CW, *et al.* RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res* 2018; **5**. DOI:10.12688/f1000research.9005.3.

517 Singh V, Kirtipal N, Song B, Lee S. Normalization of RNA-Seq data using adaptive trimmed mean with multi-reference. *Brief Bioinform* 2024; **25**. DOI:10.1093/bib/bbae241.

518 Stelzer G, Rosen N, Plaschkes I, *et al.* The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics* 2016; **2016**. DOI:10.1002/cpbi.5.

519 Wolf SJ, Bachtiar M, Wang J, Sim TS, Chong SS, Lee CGL. An update on ABCB1 pharmacogenetics: Insights from a 3D model into the

location and evolutionary conservation of residues corresponding to SNPs associated with drug pharmacokinetics. Pharmacogenomics Journal. 2011; **11**. DOI:10.1038/tpj.2011.16.

520 Hubert KA, Wellik DM. Hox genes in development and beyond. *Development* 2023; **150**. DOI:10.1242/dev.192476.

521 Li H ning, Li X rui, Lv Z tao, Cai M miao, Wang G, Yang Z fang. Elevated expression of FREM1 in breast cancer indicates favorable prognosis and high-level immune infiltration status. *Cancer Med* 2020; **9**. DOI:10.1002/cam4.3543.

522 Coutelier M, Jacoupy M, Janer A, *et al.* NPTX1 mutations trigger endoplasmic reticulum stress and cause autosomal dominant cerebellar ataxia. *Brain* 2022; **145**. DOI:10.1093/brain/awab407.

523 Zhao Y, Yu Y, Zhao W, *et al.* As a downstream target of the AKT pathway, NPTX1 inhibits proliferation and promotes apoptosis in hepatocellular carcinoma. *Biosci Rep* 2019; **39**. DOI:10.1042/BSR20181662.

524 van der Vaart B, Franker MAM, Kuijpers M, *et al.* Microtubule plus-end tracking proteins SLAIN1/2 and ch-TOG promote axonal development. *Journal of Neuroscience* 2012; **32**. DOI:10.1523/JNEUROSCI.1240-12.2012.

525 Zhen Y, Pavez M, Li X. The role of Pcdh10 in neurological disease and cancer. J. Cancer Res. Clin. Oncol. 2023; **149**. DOI:10.1007/s00432-023-04743-w.

526 Bianchi ME, Mezzapelle R. The Chemokine Receptor CXCR4 in Cell Proliferation and Tissue Regeneration. Front. Immunol. 2020; **11**. DOI:10.3389/fimmu.2020.02109.

527 Palmada M, Schmalisch K, Böhmer C, *et al.* Loss of function mutations of the GJB2 gene detected in patients with DFNB1-associated hearing impairment. *Neurobiol Dis* 2006; **22**. DOI:10.1016/j.nbd.2005.10.005.

528 Zheng H, Qiao C, Tang R, *et al.* Follistatin N terminus differentially regulates muscle size and fat In Vivo. *Exp Mol Med* 2017; **49**. DOI:10.1038/emm.2017.135.

529 Wu G, Ma Z, Cheng Y, *et al.* Targeting Gas6/TAM in cancer cells and tumor microenvironment. Mol. Cancer. 2018; **17**. DOI:10.1186/s12943-018-0769-1.

530 Tews D, Fromme T, Keuper M, *et al.* Teneurin-2 (TENM2) deficiency induces UCP1 expression in differentiating human fat cells. *Mol Cell Endocrinol* 2017; **443**. DOI:10.1016/j.mce.2017.01.015.

531 Ning WR, Jiang D, Liu XC, *et al.* Carbonic anhydrase XII mediates the survival and prometastatic functions of macrophages in human hepatocellular carcinoma. *Journal of Clinical Investigation* 2022; **132**. DOI:10.1172/JCI153110.

532 Fakiola M, Singh OP, Syn G, *et al.* Transcriptional blood signatures for active and amphotericin B treated visceral leishmaniasis in India. *PLoS Negl Trop Dis* 2019; **13**. DOI:10.1371/journal.pntd.0007673.

533 Scott TM, Solis-Leal A, Lopez JB, Robison RA, Berges BK, Pickett BE. Comparison of Intracellular Transcriptional Response of NHBE Cells to Infection with SARS-CoV-2 Washington and New York Strains. *Front Cell Infect Microbiol* 2022; **12**. DOI:10.3389/fcimb.2022.1009328.

534 Guo D, Zhang M, Wei T, *et al.* NFKBIZ regulates NFκB signaling pathway to mediate tumorigenesis and metastasis of hepatocellular carcinoma by direct interaction with TRIM16. *Cell Mol Life Sci* 2024; **81**: 167.

535 Perie L, Verma N, Mueller E. The forkhead box transcription factor FoxP4 regulates thermogenic programs in adipocytes. *J Lipid Res* 2021; **62**. DOI:10.1016/J.JLR.2021.100102.

536 Kaur P, Gupta M, Sagar V. FUT2 gene as a genetic susceptible marker of infectious diseases: A Review. *Int J Mol Epidemiol Genet* 2022; **13**.

537 Dal Col J, Lamberti MJ, Nigro A, *et al.* Phospholipid scramblase 1: a protein with multiple functions via multiple molecular interactors. Cell Communication and Signaling. 2022; **20**. DOI:10.1186/s12964-022-00895-3.

538 Suzuki H, Yoshida T, Morisada N, *et al.* De novo NSF mutations cause early infantile epileptic encephalopathy. *Ann Clin Transl Neurol* 2019; **6**. DOI:10.1002/acn3.50917.

539 McEwan DG, Popovic D, Gubas A, *et al.* PLEKHM1 regulates autophagosome-lysosome fusion through HOPS complex and LC3/GABARAP proteins. *Mol Cell* 2015; **57**. DOI:10.1016/j.molcel.2014.11.006.

540 Szymura SJ, Bernal GM, Wu L, *et al.* DDX39B interacts with the pattern recognition receptor pathway to inhibit NF-κB and sensitize to alkylating chemotherapy. *BMC Biol* 2020; **18**. DOI:10.1186/s12915-020-0764-z.

541 Sun P, Zhang H, Shi J, *et al.* KRTCAP2 as an immunological and prognostic biomarker of hepatocellular carcinoma. *Colloids Surf B Biointerfaces* 2023; **222**. DOI:10.1016/j.colsurfb.2023.113124.

542 Li J, Xue L, Wang J, *et al.* Activation of the Chemokine Receptor CCR1 and Preferential Recruitment of Gαi Suppress RSV Replication: Implications for Developing Novel Respiratory Syncytial Virus Treatment Strategies. *J Virol* 2022; **96**. DOI:10.1128/jvi.01309-22.

543 Liu J-C, Pan Z-N, Ju J-Q, *et al.* Kinesin KIF3A regulates meiotic progression and spindle assembly in oocyte meiosis. *Cell Mol Life Sci* 2024; **81**: 168.

544 Li HT, Wei B, Li ZQ, *et al.* Diagnostic and prognostic value of MCM3 and its interacting proteins in hepatocellular carcinoma. *Oncol Lett* 2020; **20**. DOI:10.3892/OL.2020.12171.

545 Van De Mark D, Kong D, Loncarek J, Stearns T. MDM1 is a microtubule-binding protein that negatively regulates centriole duplication. *Mol Biol Cell* 2015; **26**. DOI:10.1091/mbc.E15-04-0235.

546 Kong C, Zhu Y, Xie X, Wu J, Qian M. Six potential biomarkers in septic shock: a deep bioinformatics and prospective observational study. *Front Immunol* 2023; **14**. DOI:10.3389/fimmu.2023.1184700.

547 Fu Z, Runquist JA, Montgomery C, Miziorko HM, Kim JJP. Functional insights into human HMG-CoA lyase from structures of acyl-CoA-containing ternary complexes. *Journal of Biological Chemistry* 2010; **285**. DOI:10.1074/jbc.M110.139931.

548 Ammons MCB, Siemsen DW, Nelson-Overton LK, Quinn MT, Gauss KA. Binding of pleomorphic adenoma gene-like 2 to the tumor necrosis factor (TNF)-α-responsive region of the NCF2 promoter regulates p67 phox expression and NADPH oxidase activity. *Journal of Biological Chemistry* 2007; **282**. DOI:10.1074/jbc.M610618200.

549 Li R, Zhao M, Sun M, *et al.* P4HA2 promotes the progression of papillary thyroid cancer by enhancing degradation of IκBα to activate NF-κB signaling pathway. *Cancer Cell Int* 2025; **25**: 249.

550 Iwanami N, Higuchi T, Sasano Y, *et al.* WDR55 is a nucleolar modulator of ribosomal RNA synthesis, cell cycle progression, and teleost organ development. *PLoS Genet* 2008; **4**. DOI:10.1371/journal.pgen.1000171.

551 Khanaliha K, Sadri Nahand J, Khatami AR, *et al.* Analyzing the expression pattern of the noncoding RNAs (HOTAIR, PVT-1, XIST, H19, and miRNA-34a) in PBMC samples of patients with COVID-19, according to the disease severity in Iran during 2022–2023: A cross-sectional study. *Health Sci Rep* 2024; **7**. DOI:10.1002/hsr2.1861.

552 Tan YR, Peng D, Chen CM, Qin XQ. Nonstructural protein-1 of respiratory syncytial virus regulates HOX gene expression through interacting with histone. *Mol Biol Rep* 2013; **40**. DOI:10.1007/s11033-012-2107-9.

553 Chen S, Tang W, Yu G, Tang Z, Liu E. CXCL12/CXCR4 Axis is Involved in the Recruitment of NK Cells by HMGB1 Contributing to Persistent

Airway Inflammation and AHR During the Late Stage of RSV Infection. *Journal of Microbiology* 2023; **61**. DOI:10.1007/s12275-023-00018-8.

554 Brasch-Andersen C, Christiansen L, Tan Q, Haagerup A, Vestbo J, Kruse TA. Possible gene dosage effect of glutathione-S-transferases on atopic asthma: Using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers. *Hum Mutat* 2004; **24**. DOI:10.1002/humu.20074.

555 van de Wetering C, Elko E, Berg M, *et al.* Glutathione S-transferases and their implications in the lung diseases asthma and chronic obstructive pulmonary disease: Early life susceptibility? *Redox Biol* 2021; **43**. DOI:10.1016/j.redox.2021.101995.

556 Jones KL, Mansell A, Patella S, *et al.* Activin A is a critical component of the inflammatory response, and its binding protein, follistatin, reduces mortality in endotoxemia. *Proc Natl Acad Sci U S A* 2007; **104**. DOI:10.1073/pnas.0705971104.

557 Kariyawasam HH, Pegorier S, Barkans J, *et al.* Activin and transforming growth factor-β signaling pathways are activated after allergen challenge in mild asthma. *Journal of Allergy and Clinical Immunology* 2009; **124**. DOI:10.1016/j.jaci.2009.06.022.

558 Barnum KJ, O'Connell MJ. Cell cycle regulation by checkpoints. *Methods Mol Biol* 2014; **1170**: 29–40.

559 Wu W, Munday DC, Howell G, Platt G, Barr JN, Hiscox JA. Characterization of the Interaction between Human Respiratory Syncytial Virus and the Cell Cycle in Continuous Cell Culture and Primary Human Airway Epithelial Cells. *J Virol* 2011; **85**. DOI:10.1128/jvi.05164-11.

560 Pastey MK, Lupfer C. Host cellular transcriptional response to respiratory syncytial virus infection in HEp-2 cells: insights from cDNA microarray and quantitative PCR analyses. *Front Cell Infect Microbiol* 2025; **15**: 1613386.

561 Daynes RA, Jones DC. Emerging roles of PPARs in inflammation and immunity. Nat. Rev. Immunol. 2002; **2**. DOI:10.1038/nri912.

562 Arnold R, Neumann M, König W. Peroxisome proliferator-activated receptor-γ agonists inhibit respiratory syncytial virus-induced expression of intercellular adhesion molecule-1 in human lung epithelial cells. *Immunology* 2007; **121**. DOI:10.1111/j.1365-2567.2006.02539.x.

563 Moll R, Divo M, Langbein L. The human keratins: Biology and pathology. Histochem. Cell Biol. 2008; **129**. DOI:10.1007/s00418-008-0435-6.

564 Lacy P. Mechanisms of Degranulation in Neutrophils. *Allergy, Asthma & Clinical Immunology* 2006; **2**. DOI:10.1186/1710-1492-2-3-98.

565 Kirsebom FCM, Kausar F, Nuriev R, Makris S, Johansson C. Neutrophil recruitment and activation are differentially dependent on MyD88/TRIF and MAVS signaling during RSV infection. *Mucosal Immunol* 2019; **12**. DOI:10.1038/s41385-019-0190-0.

566 Fallon KE. The acute phase response and exercise: The ultramarathon as prototype exercise. *Clinical Journal of Sport Medicine* 2001; **11**. DOI:10.1097/00042752-200101000-00007.

567 Yanagisawa H, Davis EC. Unraveling the mechanism of elastic fiber assembly: The roles of short fibulins. International Journal of Biochemistry and Cell Biology. 2010; **42**. DOI:10.1016/j.biocel.2010.03.009.

568 Wootton SC, Kim DS, Kondoh Y, *et al.* Viral infection in acute exacerbation of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; **183**. DOI:10.1164/rccm.201010-1752OC.

569 Zhou W, Van Sinderen M, Rainczuk K, *et al.* Dysregulated miR-124-3p in endometrial epithelial cells reduces endometrial receptivity by altering polarity and adhesion. *Proc Natl Acad Sci U S A* 2024; **121**: e2401071121.

570 Sharma S, Chung CY, Uryu S, *et al.* Discovery of a highly potent, selective, orally bioavailable inhibitor of KAT6A/B histone acetyltransferases with efficacy against KAT6A-high ER+ breast cancer. *Cell Chem Biol* 2023; **30**. DOI:10.1016/j.chembiol.2023.07.005.

571 Pearson G, Robinson F, Beers Gibson T, *et al.* Mitogen-Activated Protein (MAP) Kinase Pathways: Regulation and Physiological Functions*. *Endocr Rev* 2001; **22**. DOI:10.1210/edrv.22.2.0428.

572 Dudley DT, Pang L, Decker SJ, Bridges AJ, Saltiel AR. A synthetic inhibitor of the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci U S A* 1995; **92**. DOI:10.1073/pnas.92.17.7686.

573 Zaheer RS, Koetzler R, Holden NS, Wiehler S, Proud D. Selective Transcriptional Down-Regulation of Human Rhinovirus-Induced Production of CXCL10 from Airway Epithelial Cells via the MEK1 Pathway. *The Journal of Immunology* 2009; **182**. DOI:10.4049/jimmunol.0802401.

574 Tacon CE, Newton R, Proud D, Leigh R. Rhinovirus-Induced MMP-9 Expression Is Dependent on Fra-1, Which Is Modulated by Formoterol and Dexamethasone. *The Journal of Immunology* 2012; **188**. DOI:10.4049/jimmunol.1101666.

575 Chen W, Monick MM, Carter AB, Hunninghake GW. Activation of ERK2 by respiratory syncytial virus in A549 cells is linked to the production of interleukin 8. *Exp Lung Res* 2000; **26**. DOI:10.1080/019021400269934.

576 Garcia-Vello P, Di Lorenzo F, Zucchetta D, Zamyatina A, De Castro C, Molinaro A. Lipopolysaccharide lipid A: A promising molecule for new immunity-based therapies and antibiotics. Pharmacol. Ther. 2022; **230**. DOI:10.1016/j.pharmthera.2021.107970.

577 Monick MM, Yarovinsky TO, Powers LS, *et al.* Respiratory Syncytial Virus Up-regulates TLR4 and Sensitizes Airway Epithelial Cells to Endotoxin. *Journal of Biological Chemistry* 2003; **278**. DOI:10.1074/jbc.M308093200.

578 Timmermans S, Souffriau J, Libert C. A general introduction to glucocorticoid biology. Front. Immunol. 2019; **10**. DOI:10.3389/fimmu.2019.01545.

579 Marr N, Turvey SE. Role of human TLR4 in respiratory syncytial virus-induced NF-κB activation, viral entry and replication. *Innate Immun* 2012; **18**. DOI:10.1177/1753425912444479.

580 Shao J, Youngblood H, Yang L. Targeting SPI1 to mitigate amyloid-β pathology in Alzheimer's disease. *J Alzheimers Dis* 2025; **104**: 334–7.

581 Bjune JI, Dyer L, Røsland G V., *et al.* The homeobox factor Irx3 maintains adipogenic identity. *Metabolism* 2020; **103**. DOI:10.1016/j.metabol.2019.154014.

582 Wang Q, Wang X, Li J, Yin T, Wang Y, Cheng L. PRKCSH serves as a potential immunological and prognostic biomarker in pan-cancer. *Sci Rep* 2024; **14**. DOI:10.1038/s41598-024-52153-w.

583 Zhu Z, Tang W, Ray A, *et al.* Rhinovirus stimulation of interleukin-6 in vivo and in vitro. Evidence for nuclear factor kappa B-dependent transcriptional activation. *J Clin Invest* 1996; **97**: 421–30.

584 Levitz R, Wattier R, Phillips P, *et al.* Induction of IL-6 and CCL5 (RANTES) in human respiratory epithelial (A549) cells by clinical isolates of respiratory syncytial virus is strain specific. *Virol J* 2012; **9**. DOI:10.1186/1743-422X-9-190.

585 Hornsleth A, Klug B, Nir M, *et al.* Severity of respiratory syncytial virus disease related to type and genotype of virus and to cytokine values in nasopharyngeal secretions. *Pediatric Infectious Disease Journal* 1998; **17**. DOI:10.1097/00006454-199812000-00003.

586 Yaugel-Novoa M, Bourlet T, Longet S, Botelho-Nevers E, Paul S. Association of IFNAR1 and IFNAR2 with COVID-19 severity. *Lancet Microbe* 2023; **4**: e487.

587 Bhowmick R, Gappa-Fahlenkamp H. Cells and Culture Systems Used to Model the Small Airway Epithelium. Lung. 2016; **194**. DOI:10.1007/s00408-016-9875-2.

588 Pace E, Di Vincenzo S, Ferraro M, Lanata L, Scaglione F. Role of airway epithelium in viral respiratory infections: Can carbocysteine prevent or mitigate them? Immunology. 2024; **172**. DOI:10.1111/imm.13762.

589 Segeritz CP, Vallier L. Cell Culture: Growing Cells as Model Systems In Vitro. In: Basic Science Methods for Clinical Researchers. 2017. DOI:10.1016/B978-0-12-803077-6.00009-6.

590 Michi AN, Proud D. A toolbox for studying respiratory viral infections using air-liquid interface cultures of human airway epithelial cells. *Am J Physiol Lung Cell Mol Physiol* 2021; **321**. DOI:10.1152/AJPLUNG.00141.2021.

591 Leung C, Wadsworth SJ, Jasemine Yang S, Dorscheid DR. Structural and functional variations in human bronchial epithelial cells cultured in air-liquid interface using different growth media. *Am J Physiol Lung Cell Mol Physiol* 2020; **318**. DOI:10.1152/AJPLUNG.00190.2019.

592 Bridges JP, Vladar EK, Huang H, Mason RJ. Respiratory epithelial cell responses to SARS-CoV-2 in COVID-19. Thorax. 2022; **77**. DOI:10.1136/thoraxjnl-2021-217561.

593 Karakocak BB, Keshavan S, Gunasingam G, *et al.* Rethinking of TEER measurement reporting for epithelial cells grown on permeable inserts. *European Journal of Pharmaceutical Sciences* 2023; **188**. DOI:10.1016/j.ejps.2023.106511.

594 Stahl D, Richard K, Papadimos T. Complications of bronchoscopy: A concise synopsis. *Int J Crit Illn Inj Sci* 2015; **5**. DOI:10.4103/2229-5151.164995.

595 Calvert BA, Ryan AL. Application of iPSC to Modelling of Respiratory Diseases. In: Advances in Experimental Medicine and Biology. 2020. DOI:10.1007/5584_2019_430.

596 Inamori S, Fujii M, Satake S, *et al.* Modeling early stages of endoderm development in epiblast stem cell aggregates with supply of extracellular matrices. *Dev Growth Differ* 2020; **62**. DOI:10.1111/dgd.12663.

597 Solnica-Krezel L, Sepich DS. Gastrulation: Making and shaping germ layers. Annu. Rev. Cell Dev. Biol. 2012; **28**. DOI:10.1146/annurev-cellbio-092910-154043.

598 Inamori S, Fujii M, Satake S, *et al.* Modeling early stages of endoderm development in epiblast stem cell aggregates with supply of extracellular matrices. *Dev Growth Differ* 2020; **62**. DOI:10.1111/dgd.12663.

599 Wang P, Rodriguez RT, Wang J, Ghodasara A, Kim SK. Targeting SOX17 in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. *Cell Stem Cell* 2011; **8**. DOI:10.1016/j.stem.2011.01.017.

600 Wang P, Rodriguez RT, Wang J, Ghodasara A, Kim SK. Targeting SOX17 in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. *Cell Stem Cell* 2011; **8**. DOI:10.1016/j.stem.2011.01.017.

601 Kearns NA, Genga RMJ, Ziller M, *et al.* Generation of organized anterior foregut epithelia from pluripotent stem cells using small molecules. *Stem Cell Res* 2013; **11**. DOI:10.1016/j.scr.2013.06.007.

602 Green MD, Chen A, Nostro MC, *et al.* Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nat Biotechnol* 2011; **29**. DOI:10.1038/nbt.1788.

603 Minoo P. Transcriptional regulation of lung development: Emergence of specificity. Respir. Res. 2000; **1**. DOI:10.1186/rr20.

604 Park KS, Wells JM, Zorn AM, Wert SE, Whitsett JA. Sox17 influences the differentiation of respiratory epithelial cells. *Dev Biol* 2006; **294**. DOI:10.1016/j.ydbio.2006.02.038.

605 Cardoso W V., Itoh A, Nogawa H, Mason I, Brody JS. FGF-1 and FGF-7 induce distinct patterns of growth and differentiation in embryonic lung epithelium. *Developmental Dynamics* 1997; **208**. DOI:10.1002/(SICI)1097-0177(199703)208:3<398::AID-AJA10>3.0.CO;2-X.

606 Longmire TA, Ikonomou L, Hawkins F, *et al.* Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* 2012; **10**. DOI:10.1016/j.stem.2012.01.019.

607 Mou H, Zhao R, Sherwood R, *et al.* Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs. *Cell Stem Cell* 2012; **10**. DOI:10.1016/j.stem.2012.01.018.

608 Huang SXL, Islam MN, O'Neill J, *et al.* Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nat Biotechnol* 2014; **32**. DOI:10.1038/nbt.2754.

609 Firth AL, Dargitz CT, Qualls SJ, *et al.* Generation of multiciliated cells in functional airway epithelia from human induced pluripotent stem cells. *Proc Natl Acad Sci U S A* 2014; **111**. DOI:10.1073/pnas.1403470111.

610 Wong AP, Chin S, Xia S, Garner J, Bear CE, Rossant J. Efficient generation of functional CFTR-expressing airway epithelial cells from human pluripotent stem cells. *Nat Protoc* 2015; **10**. DOI:10.1038/nprot.2015.021.

611 McCauley KB, Hawkins F, Serra M, Thomas DC, Jacob A, Kotton DN. Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. *Cell Stem Cell* 2017; **20**. DOI:10.1016/j.stem.2017.03.001.

612 Ghaedi M, Le A V., Hatachi G, *et al.* Bioengineered lungs generated from human iPSCs-derived epithelial cells on native extracellular matrix. *J Tissue Eng Regen Med* 2018; **12**. DOI:10.1002/term.2589.

613 Kesimer M, Kirkham S, Pickles RJ, *et al.* Tracheobronchial air-liquid interface cell culture: A model for innate mucosal defense of the upper

airways? *Am J Physiol Lung Cell Mol Physiol* 2009; **296**. DOI:10.1152/ajplung.90388.2008.

614 Pezzulo AA, Starner TD, Scheetz TE, *et al.* The air-liquid interface and use of primary cell cultures are important to recapitulate the transcriptional profile of in vivo airway epithelia. *Am J Physiol Lung Cell Mol Physiol* 2011; **300**. DOI:10.1152/ajplung.00256.2010.

615 Dvorak A, Tilley AE, Shaykhiev R, Wang R, Crystal RG. Do airway epithelium air-liquid cultures represent the in vivo airway epithelium transcriptome? *Am J Respir Cell Mol Biol* 2011; **44**. DOI:10.1165/rcmb.2009-0453OC.

616 Pezzulo AA, Starner TD, Scheetz TE, *et al.* The air-liquid interface and use of primary cell cultures are important to recapitulate the transcriptional profile of in vivo airway epithelia. *Am J Physiol Lung Cell Mol Physiol* 2011; **300**. DOI:10.1152/ajplung.00256.2010.

617 Ji J, Ganguly K, Mihai X, *et al.* Exposure of normal and chronic bronchitis-like mucosa models to aerosolized carbon nanoparticles: comparison of pro-inflammatory oxidative stress and tissue injury/repair responses. *Nanotoxicology* 2019; **13**. DOI:10.1080/17435390.2019.1655600.

618 Stahl D, Richard K, Papadimos T. Complications of bronchoscopy: A concise synopsis. *Int J Crit Illn Inj Sci* 2015; **5**. DOI:10.4103/2229-5151.164995.

619 Johnson SM, McNally BA, Ioannidis I, *et al.* Respiratory Syncytial Virus Uses CX3CR1 as a Receptor on Primary Human Airway Epithelial Cultures. *PLoS Pathog* 2015; **11**. DOI:10.1371/journal.ppat.1005318.

620 Hallak LK, Collins PL, Knudson W, Peeples ME. Iduronic acid-containing glycosaminoglycans on target cells are required for efficient respiratory syncytial virus infection. *Virology* 2000; **271**. DOI:10.1006/viro.2000.0293.

621 Krusat T, Streckert HJ. Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch Virol* 1997; **142**. DOI:10.1007/s007050050156.

622 Dalskov L, Gad HH, Hartmann R. Viral recognition and the antiviral interferon response. *EMBO J* 2023; **42**. DOI:10.15252/embj.2022112907.

623 Haddad F, Baldwin KM. Reverse transcription of the ribonucleic acid: the first step in RT-PCR assay. *Methods Mol Biol* 2010; **630**. DOI:10.1007/978-1-60761-629-0_17.

624 Srinivasan B, Kolli AR, Esch MB, Abaci HE, Shuler ML, Hickman JJ. TEER Measurement Techniques for In Vitro Barrier Model Systems. J. Lab. Autom. 2015; **20**. DOI:10.1177/2211068214561025.

625 Dymond JS. Chapter Twenty Three - Explanatory Chapter: Quantitative PCR. In: Methods in enzymology. 2013.

626 Bas A, Forsberg G, Hammarström S, Hammarström ML. Utility of the housekeeping genes 18S rRNA, β-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand J Immunol* 2004; **59**. DOI:10.1111/j.0300-9475.2004.01440.x.

627 Mosqueira D, Mannhardt I, Bhagwan JR, *et al.* CRISPR/Cas9 editing in human pluripotent stemcell-cardiomyocytes highlights arrhythmias, hypocontractility, and energy depletion as potential therapeutic targets for hypertrophic cardiomyopathy. *Eur Heart J* 2018; **39**. DOI:10.1093/eurheartj/ehy249.

628 Kleinman HK, Martin GR. Matrigel: Basement membrane matrix with biological activity. Semin. Cancer Biol. 2005; **15**. DOI:10.1016/j.semcancer.2005.05.004.

629 Passaniti A, Kleinman HK, Martin GR. Matrigel: history/background, uses, and future applications. J. Cell Commun. Signal. 2022; **16**. DOI:10.1007/s12079-021-00643-1.

630 Ricardo R, Phelan K. Trypsinizing and subculturing mammalian cells. *Journal of Visualized Experiments* 2008. DOI:10.3791/755.

631 Tsuji K, Ojima M, Otabe K, *et al.* Effects of different cell-detaching methods on the viability and cell surface antigen expression of synovial mesenchymal stem cells. *Cell Transplant* 2017; **26**. DOI:10.3727/096368917X694831.

632 Serls AE, Doherty S, Parvatiyar P, Wells JM, Deutsch GH. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development* 2005; **132**. DOI:10.1242/dev.01570.

633 Bellusci S, Furuta Y, Rush MG, Henderson R, Winnier G, Hogan BLM. Involvement of Sonic hedgehog (Shh) in mouse embryonic lung growth and morphogenesis. *Development* 1997; **124**. DOI:10.1242/dev.124.1.53.

634 Kim SK, Melton DA. Pancreas development is promoted by cyclopamine, a hedgehog signaling inhibitor. *Proc Natl Acad Sci U S A* 1998; **95**. DOI:10.1073/pnas.95.22.13036.

635 Yuan T, Volckaert T, Chanda D, Thannickal VJ, de Langhe SP. Fgf10 Signaling in Lung Development, Homeostasis, Disease, and Repair After Injury. *Front Genet* 2018; **9**. DOI:10.3389/fgene.2018.00418.

636 Danopoulos S, Belgacemi R, Hein RFC, *et al.* FGF18 promotes human lung branching morphogenesis through regulating mesenchymal progenitor cells. *Am J Physiol Lung Cell Mol Physiol* 2023; **324**. DOI:10.1152/ajplung.00316.2022.

637 Molloy E, Adams A, Moore JB, *et al.* BMP4 induces an epithelial-mesenchymal transition-like response in adult airway epithelial cells. *Growth Factors* 2008; **26**. DOI:10.1080/08977190801987166.

638 Aziz SJ, Zeman-Pocrnich CE. Tissue Processing. In: Methods in Molecular Biology. 2022. DOI:10.1007/978-1-0716-1948-3_4.

639 McKinnon KM. Flow cytometry: An overview. *Curr Protoc Immunol* 2018; **2018**. DOI:10.1002/cpim.40.

640 McCall MN, McMurray HR, Land H, Almudevar A. On non-detects in qPCR data. *Bioinformatics* 2014; **30**. DOI:10.1093/bioinformatics/btu239.

641 Bai J, Smock SL, Jackson GR, *et al.* Phenotypic responses of differentiated asthmatic human airway epithelial cultures to rhinovirus. *PLoS One* 2015; **10**. DOI:10.1371/journal.pone.0118286.

642 Anderson CS, Chirkova T, Slaunwhite CG, *et al.* CX3CR1 Engagement by Respiratory Syncytial Virus Leads to Induction of Nucleolin and Dysregulation of Cilium-Related Genes. *J Virol* 2021; **95**. DOI:10.1128/jvi.00095-21.

643 Vaid R, Mendez A, Thombare K, *et al.* Global loss of cellular m6A RNA methylation following infection with different SARS-CoV-2 variants. *Genome Res* 2023; **33**. DOI:10.1101/gr.276407.121.

644 Zhu J, Ghosh A, Sarkar SN. OASL - A new player in controlling antiviral innate immunity. Curr. Opin. Virol. 2015; **12**. DOI:10.1016/j.coviro.2015.01.010.

645 Maelfait J, Rehwinkel J. The Z-nucleic acid sensor ZBP1 in health and disease. Journal of Experimental Medicine. 2023; **220**. DOI:10.1084/JEM.20221156.

646 Bidgood GM, Keating N, Doggett K, Nicholson SE. SOCS1 is a critical checkpoint in immune homeostasis, inflammation and tumor immunity. Front. Immunol. 2024; **15**. DOI:10.3389/fimmu.2024.1419951.

647 Göttig L, Weiß C, Stubbe M, *et al.* Apobec3A Deamination Functions Are Involved in Antagonizing Efficient Human Adenovirus Replication and Gene Expression. *mBio* 2023; **14**. DOI:10.1128/mbio.03478-22.

648 Villamayor L, López-García D, Rivero V, Martínez-Sobrido L, Nogales A, DeDiego ML. The IFN-stimulated gene IFI27 counteracts innate immune responses after viral infections by interfering with RIG-I signaling. *Front Microbiol* 2023; **14**. DOI:10.3389/fmicb.2023.1176177.

649 Ikegami K, Mukai M, Tsuchida JI, Heier RL, MacGregor GR, Setou M. TTLL7 is a mammalian β-tubulin polyglutamylase required for growth of MAP2-positive neurites. *Journal of Biological Chemistry* 2006; **281**. DOI:10.1074/jbc.M603984200.

650 Martino J, Liu Q, Vukojevic K, *et al.* Mouse and human studies support DSTYK loss of function as a low-penetrance and variable expressivity risk factor for congenital urinary tract anomalies. *Genetics in Medicine* 2023; **25**. DOI:10.1016/j.gim.2023.100983.

651 Yin B, Li MJ, Sun JL, *et al.* Novel susceptibility gene SLC23A2 functions via PI3K-AKT-mTOR pathway in etiology of non-syndromic cleft palate. *J Hum Genet* 2025; **70**. DOI:10.1038/s10038-025-01352-y.

652 Wang S, Sun Y, Li C, *et al.* TH1L involvement in colorectal cancer pathogenesis by regulation of CCL20 through the NF-κB signalling pathway. *J Cell Mol Med* 2024; **28**. DOI:10.1111/jcmm.18391.

653 Huang C, Sharma A, Thakur R, *et al.* Asporin, an extracellular matrix protein, is a beneficial regulator of cardiac remodeling. *Matrix Biology* 2022; **110**. DOI:10.1016/j.matbio.2022.04.005.

654 Bergwitz C, Roslin NM, Tieder M, *et al.* SLC34A3 mutations in patients with hereditary hypophosphatemic rickets with hypercalciuria predict a key role for the sodium-phosphate cotransporter NaPi-IIc in maintaining phosphate homeostasis. *Am J Hum Genet* 2006; **78**. DOI:10.1086/499409.

655 Li K, Markosyan RM, Zheng YM, *et al.* IFITM Proteins Restrict Viral Membrane Hemifusion. *PLoS Pathog* 2013; **9**. DOI:10.1371/journal.ppat.1003124.

656 Silverman RH. Viral Encounters with 2′,5′-Oligoadenylate Synthetase and RNase L during the Interferon Antiviral Response. *J Virol* 2007; **81**. DOI:10.1128/jvi.01471-07.

657 Lockbaum GJ, Lynes MM, Shen SA, *et al.* Characterization of exoribonuclease XRN1 as a cancer target and identification of

adenosine-3',5'-bisphosphate as a potent enzyme inhibitor. *Commun Biol* 2025; **8**. DOI:10.1038/s42003-025-08005-y.

658 Harrington J, Wheway G, Willaime-Morawek S, Gibson J, Walters ZS. Pathogenic KDM5B variants in the context of developmental disorders. Biochim. Biophys. Acta Gene Regul. Mech. 2022; **1865**. DOI:10.1016/j.bbagrm.2022.194848.

659 Liebl MC, Moehlenbrink J, Becker H, *et al.* DAZAP2 acts as specifier of the p53 response to DNA damage. *Nucleic Acids Res* 2021; **49**. DOI:10.1093/nar/gkab084.

660 Hernandez I, Botana L, Diez-Mata J, *et al.* CAP1: a novel extracellular vesicle marker linked to endothelial senescence in atherosclerosis. *Biol Direct* 2025; **20**. DOI:10.1186/s13062-025-00646-7.

661 Mitsutake S, Kim TJ, Inagaki Y, Kato M, Yamashita T, Igarashi Y. Ceramide Kinase Is a Mediator of Calcium-dependent Degranulation in Mast Cells. *Journal of Biological Chemistry* 2004; **279**. DOI:10.1074/jbc.M312885200.

662 Lyck R, Enzmann G. The physiological roles of ICAM-1 and ICAM-2 in neutrophil migration into tissues. Curr. Opin. Hematol. 2015; **22**. DOI:10.1097/MOH.0000000000000103.

663 Wawrzyniak JA, Bianchi-Smiraglia A, Bshara W, *et al.* A Purine Nucleotide Biosynthesis Enzyme Guanosine Monophosphate Reductase Is a Suppressor of Melanoma Invasion. *Cell Rep* 2013; **5**. DOI:10.1016/j.celrep.2013.09.015.

664 Li J, Liu Q, Yang J, *et al.* Biological functions and clinical implications of the CMPK2 across multisystemic diseases. Cell Biosci. 2025; **15**. DOI:10.1186/s13578-025-01466-y.

665 Florez MA, Thatavarty A, Le DT, *et al.* BST2 facilitates activation of hematopoietic stem cells through ERK signaling. *Exp Hematol* 2024; **140**. DOI:10.1016/j.exphem.2024.104653.

666 Niu Y, Qian Q, Li J, *et al.* De novo variants in AGO1 recapitulate a heterogeneous neurodevelopmental disorder phenotype. *Clin Genet* 2022; **101**. DOI:10.1111/cge.14114.

667 Xiao Z, Zhang Y, She Y, Yuan G, Yang G. IPO7 promotes lipopolysaccharide-induced inflammatory responses in human dental pulp cells via p38 MAPK and NF-κB signaling pathways. *Mol Immunol* 2023; **163**. DOI:10.1016/j.molimm.2023.09.011.

668 Bukhari SA, Yasmin A, Zahoor MA, Mustafa G, Sarfraz I, Rasul A. Secreted frizzled-related protein 4 and its implication in obesity and type-2 diabetes. IUBMB Life. 2019; **71**. DOI:10.1002/iub.2123.

669 Scheibner K, Schirge S, Burtscher I, *et al.* Epithelial cell plasticity drives endoderm formation during gastrulation. *Nat Cell Biol* 2021; **23**. DOI:10.1038/s41556-021-00694-x.

670 Liu N, Wang A, Xue M, Zhu X, Liu Y, Chen M. FOXA1 and FOXA2: the regulatory mechanisms and therapeutic implications in cancer. *Cell Death Discov* 2024; **10**: 172.

671 Trinh LT, Osipovich AB, Liu B, *et al.* Single-Cell RNA Sequencing of Sox17-Expressing Lineages Reveals Distinct Gene Regulatory Networks and Dynamic DevelopmentalTrajectories. *Stem Cells* 2023; **41**. DOI:10.1093/stmcls/sxad030.

672 Guan L, Zhao X, Tang L, *et al.* Thyroid Transcription Factor-1: Structure, Expression, Function and Its Relationship with Disease. Biomed Res. Int. 2021; **2021**. DOI:10.1155/2021/9957209.

673 Yamamoto T, Arita M, Kuroda H, Suzuki T, Kawamata S. Improving the differentiation potential of pluripotent stem cells by optimizing culture conditions. *Sci Rep* 2022; **12**. DOI:10.1038/s41598-022-18400-8.

674 Tran DN, Smith SABC, Brown DA, *et al.* Polychromatic flow cytometry is more sensitive than microscopy in detecting small monoclonal plasma cell populations. *Cytometry B Clin Cytom* 2017; **92**. DOI:10.1002/cyto.b.21401.

675 Bolling MC, Lemmink HH, Jansen GHL, Jonkman MF. Mutations in KRT5 and KRT14 cause epidermolysis bullosa simplex in 75% of the patients. *British Journal of Dermatology* 2011; **164**. DOI:10.1111/j.1365-2133.2010.10146.x.

676 Lorsuwannarat N, Kaewsanit A, Charoenpitakchai M, Ruangpratheep C, Arnutti P, Nimmanon T. Optimizing Re-staining Techniques for the Restoration of Faded Hematoxylin and Eosin-stained Histopathology Slides: A Comparative Study. *J Histochem Cytochem* 2024; **72**: 733–42.

677 Dong W, Kameyama A. Succinylation-Alcian Blue Staining of Mucins on Polyvinylidene Difluoride Membrane. In: Methods in Molecular Biology. 2024. DOI:10.1007/978-1-0716-3670-1_9.

678 Kim J, Chu J, Shen X, Wang J, Orkin SH. An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell* 2008; **132**. DOI:10.1016/j.cell.2008.02.039.

679 Ghosh S, Nehme R, Barrett LE. Greater genetic diversity is needed in human pluripotent stem cell models. Nat. Commun. 2022; **13**. DOI:10.1038/s41467-022-34940-z.

680 Carcamo-Orive I, Hoffman GE, Cundiff P, *et al.* Analysis of Transcriptional Variability in a Large Human iPSC Library Reveals Genetic and Non-genetic Determinants of Heterogeneity. *Cell Stem Cell* 2017; **20**. DOI:10.1016/j.stem.2016.11.005.

681 Ortmann D, Brown S, Czechanski A, *et al.* Naive Pluripotent Stem Cells Exhibit Phenotypic Variability that Is Driven by Genetic Variation. *Cell Stem Cell* 2020; **27**. DOI:10.1016/j.stem.2020.07.019.

682 Pantazis CB, Yang A, Lara E, *et al.* A reference human induced pluripotent stem cell line for large-scale collaborative studies. *Cell Stem Cell* 2022; **29**. DOI:10.1016/j.stem.2022.11.004.

683 Goritzka M, Durant LR, Pereira C, Salek-Ardakani S, Openshaw PJM, Johansson C. Alpha/Beta Interferon Receptor Signaling Amplifies Early

Proinflammatory Cytokine Production in the Lung during Respiratory Syncytial Virus Infection. *J Virol* 2014; **88**. DOI:10.1128/jvi.00333-14.

684 Jeong K Il, Piepenhagen PA, Kishko M, *et al.* CX3CR1 is expressed in differentiated human ciliated airway cells and co-localizes with respiratory syncytial virus on cilia in a G protein-dependent manner. *PLoS One* 2015; **10**. DOI:10.1371/journal.pone.0130517.

685 Green G, Johnson SM, Costello H, *et al.* CX3CR1 Is a Receptor for Human Respiratory Syncytial Virus in Cotton Rats. *J Virol* 2021; **95**. DOI:10.1128/jvi.00010-21.

686 Rosas-Salazar C, Chirkova T, Gebretsadik T, *et al.* Respiratory syncytial virus infection during infancy and asthma during childhood in the USA (INSPIRE): a population-based, prospective birth cohort study. *The Lancet* 2023; **401**. DOI:10.1016/S0140-6736(23)00811-5.

687 Zhou P, Cowled C, Wang LF, Baker ML. Bat Mx1 and Oas1, but not Pkr are highly induced by bat interferon and viral infection. *Dev Comp Immunol* 2013; **40**. DOI:10.1016/j.dci.2013.03.006.

688 Harioudh MK, Perez J, Chong Z, *et al.* Oligoadenylate synthetase 1 displays dual antiviral mechanisms in driving translational shutdown and protecting interferon production. *Immunity* 2024; **57**. DOI:10.1016/j.immuni.2024.02.002.

689 Roet JEG, Mikula AM, de Kok M, *et al.* Unbiased method for spectral analysis of cells with great diversity of autofluorescence spectra. *Cytometry A* 2024; **105**: 595–606.

690 Volpato V, Webber C. Addressing variability in iPSC-derived models of human disease: Guidelines to promote reproducibility. DMM Disease Models and Mechanisms. 2020; **13**. DOI:10.1242/dmm.042317.

691 Guhr A, Kobold S, Seltmann S, Seiler Wulczyn AEM, Kurtz A, Löser P. Recent Trends in Research with Human Pluripotent Stem Cells: Impact of Research and Use of Cell Lines in Experimental Research and Clinical Trials. *Stem Cell Reports* 2018; **11**. DOI:10.1016/j.stemcr.2018.06.012.

692 Popp B, Krumbiegel M, Grosch J, *et al.* Need for high-resolution Genetic Analysis in iPSC: Results and Lessons from the ForIPS Consortium. *Sci Rep* 2018; **8**. DOI:10.1038/s41598-018-35506-0.

693 Huang X, Wu SM. Isolation and functional characterization of pluripotent stem cell-derived cardiac progenitor cells. *Curr Protoc Stem Cell Biol* 2010; **Chapter 1**: Unit 1F.10.

694 Tsujisaka Y, Hatani T, Okubo C, *et al.* Purification of human iPSC-derived cells at large scale using microRNA switch and magnetic-activated cell sorting. *Stem Cell Reports* 2022; **17**. DOI:10.1016/j.stemcr.2022.05.003.

695 Carlin AF, Shresta S. Genome-wide approaches to unravelling host–virus interactions in Dengue and Zika infections. Curr. Opin. Virol. 2019; **34**. DOI:10.1016/j.coviro.2018.11.010.

696 Gomperts BN. Induction of multiciliated cells from induced pluripotent stem cells. Proc. Natl. Acad. Sci. U. S. A. 2014; **111**. DOI:10.1073/pnas.1404414111.

697 Santa Cruz A, Mendes-Frias A, Oliveira AI, *et al.* Interleukin-6 Is a Biomarker for the Development of Fatal Severe Acute Respiratory Syndrome Coronavirus 2 Pneumonia. *Front Immunol* 2021; **12**: 613422.

698 Cook D, Brown D, Alexander R, *et al.* Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. Nat. Rev. Drug Discov. 2014; **13**. DOI:10.1038/nrd4309.

699 Pushpakom S, Iorio F, Eyers PA, *et al.* Drug repurposing: Progress, challenges and recommendations. Nat. Rev. Drug Discov. 2018; **18**. DOI:10.1038/nrd.2018.168.

700 Ampuero S, Andaur R, Milano M, *et al.* Time-course of transcriptome response to respiratory syncytial virus infection in lung epithelium cells. *Acta Virol* 2018; **62**. DOI:10.4149/av_2018_225.

701 Miki K, Endo K, Takahashi S, *et al.* Efficient Detection and Purification of Cell Populations Using Synthetic MicroRNA Switches. *Cell Stem Cell* 2015; **16**. DOI:10.1016/j.stem.2015.04.005.

702 Rahman S, Ikram AR, Azeem F, Tahir Ul Qamar M, Shaheen T, Mehboob-Ur-Rahman. Precision Genome Editing with CRISPR-Cas9. *Methods Mol Biol* 2024; **2788**: 355–72.

703 Turner DL, Amoozadeh S, Baric H, Stanley E, Werder RB. Building a human lung from pluripotent stem cells to model respiratory viral infections. *Respir Res* 2024; **25**: 277.

704 Sayers EW, Beck J, Bolton EE, *et al.* Database resources of the National Center for Biotechnology Information in 2025. *Nucleic Acids Res* 2024. DOI:10.1093/nar/gkae979.