Bioinformatic Insights into the Impacts of External Stimuli on the Human Airway

by

Jennifer Ann Aguiar

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Biology

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:
Dr. Scott Tebbutt
Professor, Department of Medicine
University of British Columbia

Supervisors:
Dr. Andrew Doxey
Associate Professor, Department of Biology
University of Waterloo

Dr. Jeremy Hirota
Assistant Professor, Department of Medicine
McMaster University

Internal Members:
Dr. Brendan McConkey
Associate Professor, Department of Biology
University of Waterloo

Dr. Mungo Marsden
Associate Professor, Department of Biology
University of Waterloo

Internal-External Member:
Dr. Diane Williams
Lecturer, School of Public Health and Health Systems
University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The airway epithelium represents a critical component of the human lung that helps orchestrate defences against inhaled substances including air pollution, allergens, bacteria, and viruses. To manage these continuous insults, the airway epithelium has evolved to be a multi-functional barrier tissue with mechanical and immunological impedances that protect the lungs. In health, these coordinated functions can ensure that exposures to harmful substances are controlled to limit damage to the host, while in disease, a dysfunction in any capacity of the respiratory mucosa may lead to development or exacerbation of acute and chronic respiratory diseases. It is therefore important to understand the mechanisms that regulate respiratory mucosa function and how various stimuli may alter host defences.

Using a bioinformatic approach focused primarily on transcriptomic data, this thesis presents analyses of three prevalent inhaled particulates: tobacco smoke, cannabis smoke, and respiratory viruses. I first characterize the ATP-binding cassette (ABC) transporter gene family, a relatively unexplored contributor to respiratory mucosa biology, in the context of both tobacco smoke exposure and viral infection. I then describe my direct comparison of the effects of cannabis and tobacco smoke exposure on lung health, uncovering striking similarities relating to functional consequences, response to combination treatment, and relationship to chronic lung disease that may inform future public health policies and individual user practices. Finally, I detail how, as part of the global effort to research the novel SARS-CoV-2 virus, I obtained results suggesting an alternate model of coronavirus-host cell invasion via ACE2-independent pathways and explore the mechanisms of the host immune response induced upon viral entry, advancing the field as it grapples with the COVID-19 pandemic.

Bioinformatics coupled with large-scale data collection efforts around the world have provided a wealth of knowledge and new opportunities for biologists, including those studying respiratory medicine. Advancements in sequencing technologies and computational techniques have made possible the high-throughput analyses required to characterize the myriad changes that occur in the lungs upon exposure to various stimuli. Interdisciplinary "multi-omic" research is quickly positioning itself as a necessity for novel discoveries from within such large biological data sets. As researchers begin to adapt their experimental designs and incorporate new techniques, our collective understanding of the human airway will expand dramatically and result in increased quality of patient care and decreased global health burden. Thus, the studies presented here provide examples of how bioinformatics is essential for the advancement of the respiratory biology field.

Acknowledgements

First, I would like to thank my supervisors, Dr. Andrew Doxey and Dr. Jeremy Hirota, for their enthusiasm, encouragement, and for providing the expertise and academic freedom that has allowed me to explore numerous disciplines and passions.

I would also like to thank my committee members for their insight and advice, and the various members of the Firestone Institute for Respiratory Health and the broader Canadian respiratory community for collaborations and contributions to this thesis.

Lastly, I would like to express a great deal of gratitude to members of the Doxey and Hirota Labs, past and present. I would specifically like to thank Dr. Briallen Lobb and Dr. Michael Mansfield for their scientific discussions, guidance, friendship, and for making our windowless closet a little brighter.

Dedication

To Dalia for being my dependable confidante and a constant in my journey at UW. Thank you for being there to share in every dilemma and every success, professional and personal.

To my parents for their gentle guidance and for always being proud of me.

To my brother David and my sister Lisa. My best friends and my most avid supporters. Thank you for fostering my love of learning, teaching me how to think, and encouraging me to follow my nerdish leanings. Like so much of my work and life, it starts with you. You set the example that I strive to achieve and I am so grateful to you both.

Lastly, to my partner Will. Thank you for your patience, assurance, and unwavering love which mean more to me than I can express.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

> Lungs are a paradox. They are so fragile [...] yet, unlike our other internal organs, nestled away inside us, they are open, like a wound, to the outside world.
>
> *What Happens When You Breathe*
> Brooke Jarvis

1.1 The human airway

1.1.1 Structure and function

Almost every cell in the human body relies on oxygen to function, making the lungs and connecting airways that comprise the human respiratory system of vital importance. The primary function of the human respiratory system is to facilitate gas exchange: bringing oxygen to the body's organs while simultaneously clearing carbon dioxide waste. To achieve this, the structure of the human airway tree must fulfill various important, and often contradictory, requirements.[303] The tracheobronchial tree and terminating alveoli must take up minimal space within the chest cavity while still providing enough surface area to mediate gas exchange efficiently. There needs to be enough rigidity to provide structural integrity and maintain airflow but also enough flexibility to permit the removal of mucus and debris.[303] In fact, airway architecture is a primary determinant of how inhaled particulates deposit within the lungs, thereby having major implications to lung health and disease.

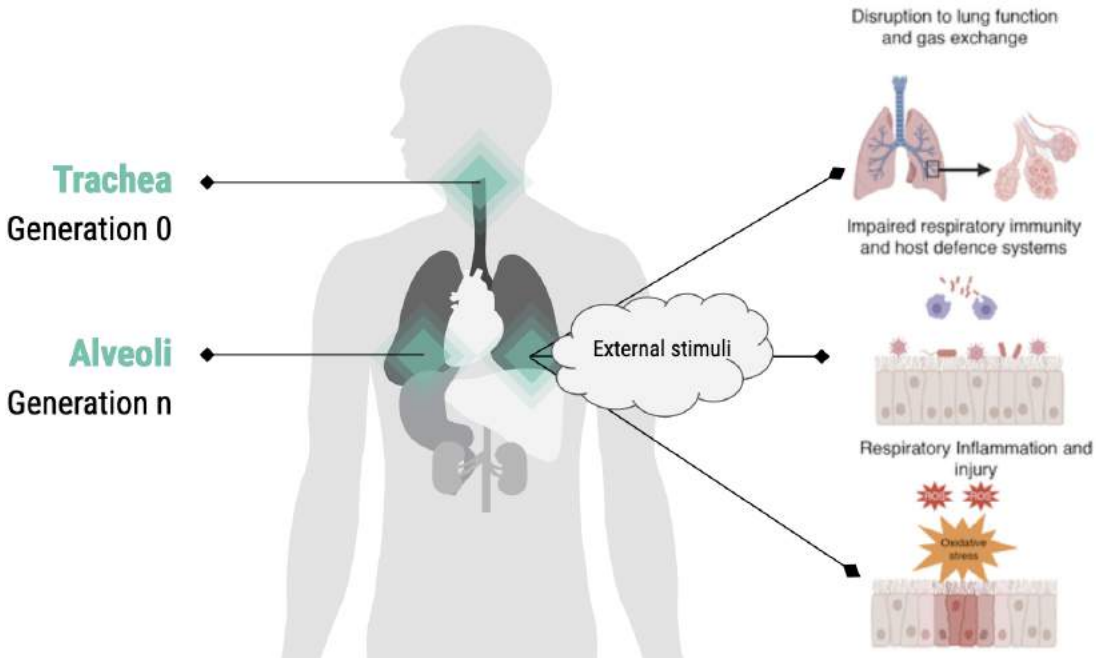


Figure 1.1: Depiction of airway tree generations. Beginning at the trachea (generation 0), the airway tree branches into bronchi, bronchioles, and alveolar ducts. Termination occurs at the alveolar sacs, which are also the site of gas exchange. Exposure to external stimuli may result in biological consequences, as depicted on the right. Figure modified from Miyashita *et al.*[262]

The human lung functions at the interface of the external and internal environments and is exposed to over 10,000 litres of air each day from normal respiration. The entire airway, in varying degrees, is therefore in frequent contact with the various particulates present in the air. This reality necessitates the well-vascularized mucus membrane composed primarily of epithelial cells, which warms and conditions the air while simultaneously cleansing and protecting the lung.[316]

1.1.2 The airway epithelium

The airway epithelium represents a critical component of the human lung that helps orchestrate defences against inhaled noxious substances that may include air pollution, allergens, bacteria, and viral insults.[164,288,310] To manage these continuous insults, the airway epithelium has evolved to be a multi-functional barrier tissue with mechanical and

immunological impedances, manifested through the mucociliary ladder, protein-protein junctions, and innate immune processes involving cytokine production. The airway epithelium extends along the entire airway tree, from the nasal cavity of the upper airways down to simple squamous epithelium.[164,288] In health, these coordinated functions can ensure that infections, allergens, and exposures to air pollution are controlled to limit damage to the host, while in disease, a dysfunction in any capacity of the respiratory mucosa may lead to development or exacerbation of chronic respiratory disease. It is therefore important to understand the mechanisms that regulate respiratory mucosa function and how prevalent airway particulates may alter host defences.

1.1.3 Airway exposure and disease

Innate immune activities of the airway epithelium rely on accurate sensing of the external environment. Exposure to various particulates can trigger the innate immune response, causing increased airway inflammation, epithelial cell damage and changes to airway expression profiles when not resolved in a timely manner. These changes can lead to various lung diseases and aberrant lung function, both acute and chronic in nature, and may be dependent on the type of stimuli to which the airway is exposed.

Tobacco smoke

An important example of an airway pollutant to which the respiratory mucosa may be exposed is tobacco cigarette smoke, which has a well characterized relationship to various lung diseases such as lung cancer and chronic obstructive pulmonary disease (COPD) via both first- and second-hand smoke exposure.

The causal link between tobacco smoke exposure and the onset of lung cancer had been suspected since the early 20th century and was confirmed in the 1940-50s through a combination of epidemiology, animal experiments, cellular pathology and chemical analytics (though lobbying from tobacco companies delayed wide-scale acceptance for several decades still).[309] As of 2020, 2.2 million people were newly diagnosed with lung cancer worldwide resulting in 1.8 million deaths (the most of any cancer).[280] Of all lung cancer cases, 80-90% are estimated to be caused by tobacco smoke exposure.[4]

Extensive evidence now exists that exposure to tobacco smoke, as well as biomass smoke generated from dried wood and charcoal, are risk factors for the development of COPD, a disease that can include both chronic bronchitis and emphysema, and lung cancers.[137,205,223,331] In addition to the approximately 100 carcinogenic, co-carcinogenic,

or mutagenic compounds found within tobacco cigarette smoke,[331] combustion of tobacco cigarettes produces polyaromatic hydrocarbons (PAHs) such as benzo[a]pyrene which are known to be associated with the production of reactive oxygen species (ROS)[392] and increased expression of the cytochrome p4501A1 isozyme gene (*CYP1A1*) that is involved in downstream metabolism of PAHs into various carcinogenic compounds. Given the complexity of tobacco smoke's chemical composition, determining exactly which active compounds are responsible for various tobacco-associated disease outcomes has proven difficult[331] and required direct studies of individual constituents. More recent studies have begun to leverage bioinformatic methods in conjunction with molecular assays and epidemiology to identify which populations are most susceptible to disease development and which genes and pathways can be developed into suitable biomarkers of tobacco smoke exposure.[255,289,353]

Cannabis smoke

Cannabis is the second most commonly smoked substance globally next to tobacco smoke[281,282] however, unlike tobacco smoke, research regarding the effects of cannabis use on lung function and disease onset is not as well understood and is often inconsistent.

While cannabis can be consumed through various routes, combustion of dried cannabis flower or "bud" is the most common delivery method. In Canada, 90% of users identified combusted smoke inhalation as a route of delivery with 94 and 89% of users reporting combustion as their primary mode of consumption in the 2017 and 2018 Canadian Cannabis Surveys, respectively.[68,69] For this reason, and since very little information is available regarding the pulmonary effects of alternative consumption methods such as vaporization, this thesis will focus on the acute and long-term effects of cannabis smoking as compared to tobacco smoking.

Like tobacco smoking, cannabis combustion produces similar toxic compounds such as PAHs,[375] which can activate *CYP1A1* upon inhalation, as mentioned above. However, unlike with tobacco smoke, research has suggested that δ-9 tetrahydrocannabinol (THC), a cannabinoid and primary component of cannabis smoke, may act as a competitive inhibitor of CYP1A1 and therefore have a mitigating effect.[327] In fact, many of the cannabinoids present in cannabis smoke have been shown to have different biologic effects and interactions compared to those observed with nicotine and tobacco smoke which makes the direct comparison of cannabis and tobacco more complex and the acute and long-term effects of cannabis smoke exposure more difficult to predict.

Cannabis use has been shown to relieve pain and reduce lung inflammation[386] however,

it has also been linked to decreased lung immune response[267] and increased risk of chronic bronchitis.[372] In contrast to tobacco smoke, the existing evidence suggests that repeated cannabis smoke exposure results in a chronic bronchitis phenotype with little evidence of emphysema.[17,371,374,420] Furthermore, unlike tobacco and biomass exposure, which are accompanied by a dose-dependent risk for development of lung cancers, a similar relationship has not been observed for repeat cannabis users despite the presence of carcinogens in cannabis smoke.[434] It should be noted that many cannabis smokers also smoke tobacco concurrently which adds confounding factors to much of the current body of cannabis-related research.[375]

It is clear that more research, including well-controlled clinical studies, needs to be performed in this area to gain a more complete understanding of the effects of cannabis on the human airway.

Respiratory viruses

Acute upper respiratory viral infections are the most common form of infection in humans,[110] with influenza A and B viruses being the leading infectious cause of morbidity and mortality amongst the elderly and children and human rhinoviruses being the leading cause of mild upper respiratory tract infections overall.[139] Symptoms of these viruses, responsible for the "flu" and "common cold", respectively, have significant overlap including headache, nasal congestion, nasal discharge, cough and sore throat,[110] and may progress to lower respiratory tract infections in severe cases. Influenza viral infections are often differentiated by the presence of fever and loss of appetite.[110] While both of these pathogens are single stranded RNA viruses, influenza viruses, which belong to the *Orthomyxoviridae* family and can be classified as either type A, B, C, or D, are enveloped and utilize a sialylated receptor for host entry[107] whereas rhinoviruses, which are members of the family *Picornaviridae* and exist as species A, B, or C comprising approximately 160 unique types, are non-enveloped and use one of three glycoproteins to gain entrance to the host cell: intercellular adhesion molecule 1 (ICAM-1), low density lipoprotein receptor (LDLR) and cadherin-related family member 3 (CDHR3).[55] The prevalence of these viruses and their functional variation require the human respiratory system to have diverse defence mechanisms in place to handle oncoming infections.

The threat posed by viruses that infect the respiratory mucosa is countered by the airway epithelium expressing functional toll-like receptors (TLRs), nucleotide binding and leucine-rich-repeat pyrin domain containing (NLRP) receptors, and cytosolic nucleic acid sensors that are able to rapidly detect exposures and provide host defence.[30,163,164,288,310,442] Antiviral sensing mechanisms in the respiratory mucosa enable responses to influenza A,

respiratory syncytial virus, rhinovirus, and human parainfluenza virus; all single stranded RNA viruses.[168] dsDNA viruses are also relevant lung infections, with adenovirus capable of inducing influenza like symptoms in healthy subjects and associated with chronic respiratory disease exacerbations.[50,72,214,239] Like RNA viruses, adenovirus is able to infect airway epithelium followed by replication, which leads to a variety of innate immune defences able to sense viral nucleic acids and proteins.[44,48,72] Vaccinia virus is another dsDNA virus that is able to infect airway epithelium and has been explored for capacity to genetically engineer the virus for transgene delivery, vaccination strategies, and studying Variola virus infections.[18,56,100,318,397]

Within the last 20 years, three highly pathogenic human coronaviruses (hCoVs) have emerged, resulting in varying degrees of global economic and public health crisis. Coronaviruses (CoVs) refer to a class of enveloped single standed RNA viruses that can be broken down into four categories: alpha, beta, gamma, and delta.[443] Of these four genera, β-CoVs have been responsible for the three major outbreaks of the last two decades.[443] In 2003, an epidemic of Severe Acute Respiratory Syndrome (SARS)-CoV resulted in 8098 infections and 774 deaths, globally. In 2012, Middle East Respiratory Syndrome (MERS)-CoV emerged resulting in an epidemic that has killed approximately 858 people to date worldwide.[3] Both SARS-CoV and MERS-CoV have a high case fatality rate, infecting the lower airways and rapidly causing acute lung injury(ALI), acute respiratory distress syndrome (ARDS), and multi-organ failure.[443] While there have been no cases of SARS for over a decade, MERS continues to be an on-going public health concern.

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) emerged in December 2019 to cause a pandemic of coronavirus disease (COVID-19).[440] Like it's predecessors, SARS-CoV-2 causes a respiratory infection, along with ARDS in severe cases. However, pre/asymptomatic airborne transmission and high viral titre early in the course of the disease significantly increases the infectiousness of SARS-CoV-2 compared to other coronaviruses such as SARS-CoV.[394] This increased transmissibility and infectivity explain how SARS-CoV-2 has been able to cause such a pandemic and why, despite the lower case fatality rate, deaths from SARS-CoV-2 far surpass those of SARS-CoV and MERS-CoV.

The seminal report identifying the receptor for SARS-CoV used a HEK293 cell over-expression system to identify angiotensin-converting enzyme 2 (ACE2) as a receptor by co-immunoprecipitation with SARS-CoV spike (S) domain 1.[238] Like SARS-CoV, SARS-CoV-2 is suggested to enter the host cell by means of the interaction between the S protein and the ACE2 receptor.[440] However, although there is evidence that SARS-CoV-2 and SARS-CoV both utilise ACE2 as a receptor to facilitate virus entry, it is possible that differences in host entry mechanisms play a role in the large epidemiological differences between the two viruses, which may include additional unidentified receptors. Furthermore,

given the novelty of the virus, much is still not fully understood about the host response elicited by the virus upon entry.

Understanding viral entry mechanisms and exploring how the airway epithelium responds to viruses may provide insights into the transmission and pathogenesis of both common and novel respiratory viruses and identify new strategies for controlling infections and implementing vaccination strategies relevant to lung health and disease.

1.1.4 Knowledge gaps

While there has been substantial progress made in the characterization of the basic mechanisms related to lung disease in recent decades, there are still many facets of how the human lung responds to external stimuli that remain poorly understood. Clinical management of various lung diseases have also improved in this time, yet our ability to predict disease onset and progression and treat patients in a personalized manner is still inadequate.[209] These areas must be addressed to tackle the rising number of respiratory illnesses around the world.

Many lung diseases are heterogeneous in nature and by the time they are identified by current clinical diagnostic tools will often have progressed to a more severe, chronic condition.[209] Advances in molecular biotechnology have allowed for earlier detection of some lung diseases such as cystic fibrosis (CF) through the development of diagnostic biomarkers (e.g., the identification and use of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene as the key marker of CF), but to capture the complexity of lung disease phenotypes and effectively stratify and diagnose disease, larger networks of genes need to be analyzed. *CFTR* represents a clinically relevant example of an ATP-binding cassette (ABC) transporter being used as a key diagnostic marker for lung disease however, this one gene represents only a fraction of the 48 ABC transporters that remain to be more extensively studied in the respiratory mucosa for expression patterns and function in lung health and disease. Moreover, when the diversity of lung disease phenotypes are taken into account, the list of affected genes that may be indicative of disease onset and progression rapidly increases. When this is coupled with the fact that many prevalent airway stimuli are either under-researched (i.e., cannabis smoke) or completely novel (i.e., SARS-CoV-2), the size of the current knowledge gap grows further.

Advancements in bioinformatic technologies allow for the large-scale analyses required to characterize the hundreds of molecular changes that occur upon exposure to various inhaled particulates. This is crucial for defining defence mechanisms, and for identifying potential diagnostic and prognostic biomarkers in airway epithelial cells. These discoveries have the

potential to lead to rapid and personalized interventions that could reduce pathologies and exacerbations of chronic respiratory diseases.

1.2 Bioinformatics in respiratory research

1.2.1 Where are we now?

The foundations for bioinformatics were laid as early as the 1960s when scientists such as Margaret Dayoff began utilizing computer technology in the fields of biochemistry and protein analysis.[124] That said, it was not until the late 1990s-early 2000s that bioinformatics really gained a foothold through parallel advancements in both computer science and biology and the sequencing of the human genome.[124] As sequencing costs began to drop during this "dawn of the genomic era", the amount of biological data being produced and deposited into public databases increased exponentially.[124] The amount of data that exists now is truly enormous and requires sophisticated tools and techniques to analyze, re-analyze, and extract useful biological and medical information. Bioinformatics is an interdisciplinary field that provides the toolkit necessary to procure, store, and assess such data.

While there have been myriad advancements over the years in the areas of bioinformatics and respiratory research, and while bioinformatic analysis is applicable to all facets of the central dogma (DNA, RNA, proteins, and beyond), this thesis will focus on the use of transcriptomics, the study of the complete set of RNA transcripts in a given sample, to answer biologically relevant questions related to the human airway.

Many lung diseases have been associated with unique mRNA expression signatures.[209] Notably, the first distinctions between Th2-high and Th2-low asthma were based on expression levels of a panel of IL-13-induced genes detected through microarray and polymerase chain reaction (PCR) analyses.[417]

Microarray analysis of primary airway epithelial cells is a low-cost and accessible method for mRNA detection and can be used to perform differential expression analysis with the goal of identifying changes in transcript expression that may act as markers for the early onset of airway diseases. However, microarray technology limits the researcher to only evaluating genes for which probes exist on the given microarray chip and has a low signal-to-noise ratio making it difficult to detect lowly abundant transcripts.[438] The genomic era of the early 2000s gave rise to high-throughput next generation sequencing (NGS) resulting in an influx of biological "'omic" data; that is, data derived from methods such as genomics, transcriptomics, or proteomics that aim to quantify and characterize the various biological

molecules that give way to cellular structure and function. RNA-sequencing has now begun to replace previous microarray technology as the gold standard by allowing researchers to explore the entire transcriptome and gain insights into the short- and long-term effects of airway particulate exposure.

NGS is a large-scale alternative to arrays that can be used for, among other things, differential expression analysis through quantitation of transcript expression levels.[272] High-throughput RNA-sequencing (RNA-seq) necessitates the conversion of RNA into complementary cDNA which can then undergo fragmentation into short reads that are then sequenced at a specified depth and coverage.[409] For most applications, 20-30X sequencing depth is considered sufficient to recall the majority of transcripts. Sequenced, fragmented reads can be assembled *de novo* or aligned to a reference transcriptome scaffold and assembled from there.[272] After assembly, numbers of each transcript are quantitated which allows for comparison of transcript abundance and differential expression analysis between cells exposed to a pollutant or derived from a patient with lung disease and healthy cells.[409] This provides a snapshot of the cells under a given physiological condition. With respect to lung disease, lung cancer is one particular area of research that has invested heavily into NGS and bioinformatics research to pinpoint mutations resulting in biomarker discovery. As mentioned previously, tobacco smoke exposure is strongly linked to the development of lung cancer and as such, our understanding of the effects of tobacco smoke exposure have also benefited greatly from bioinformatic analysis.

While there have now been a multitude of papers citing use of NGS to study the effects of tobacco smoke on respiratory health,[255,289,353] one more recent meta-analysis of lung RNA-seq data identified 5 SNPs and 2 insertions/deletions (INDELs) in *SCGB1A1*, *SCGB3A1* and *NFKB1A* genes only present in smokers with lung cancer compared to cancer-free smokers, and healthy, non-smokers.[98] Smokers with these mutations now have an indicator for disease pathogenesis which is useful for early diagnosis of lung cancer. Biomarker discovery such as this highlights the usefulness of NGS in the respiratory space.

Since the use of NGS technology has been so successful in the study of tobacco, methods such as RNA-seq are starting to be applied to cannabis research as well. While this data is currently limited, studies comparing RNA-seq expression between cannabis-exposed, tobacco-exposed, and healthy cells both *in vitro*[345] and *in vivo*[10] are beginning to elucidate some of the similarities and differences between cannabis- and tobacco-exposed epithelial expression profiles which will hopefully provide insight into cannabis-associated respiratory outcomes.

NGS technologies have also revolutionized the study of respiratory viral infections. In 1977, the ϕX174 bacteriophage, a single-stranded DNA virus that infects *E. coli*, became the

first genome ever to be sequenced using first generation methods.[332] Since then, and with the advent of NGS, countless human-infecting viruses have been sequenced, comprising what is now known as the human virome. Viral genomes and, subsequently, viral transcriptomes are incredibly small compared to the human genome, making sequencing of respiratory viruses for detection and classification readily feasible. High-throughput transcriptomics have also greatly improved human challenge studies by allowing for the changes in thousands of transcripts within a human sample to be measured after exposure to a pathogen of interest.[38] As technologies advance further, bioinformatics positions itself as a key method for the rapid identification and mechanistic understanding of both common and novel respiratory viruses.

Though there is much debate in the bioinformatic community about best practices when it comes to sequencing (depth and coverage requirements, number of replicates), the alignment, assembly, and quantitation of genes (various pipelines and software exist for each step of the protocol), and statistical analyses, it is clear that transcriptomic NGS applications have become incredibly important for identifying biomarkers and risk factors which help with the early detection of human lung disease. It is through this technology that the field of respiratory medicine has really begun to thrive.

1.2.2 Application of multi-omic approaches

While this thesis focuses primarily on the use of transcriptomics to elucidate novel insights regarding the effects of various airway exposures, the eventual goal here is the integration of multi-omic technologies in an attempt to gain a fuller understanding of the afflictions of the human respiratory system and to usher in the age of precision respiratory medicine.

Biological systems are multi-factorial in how they are governed.[184] A complete understanding of the mechanisms involved in any given disease or exposure response will therefore require an integrated approach. The decreased cost and increased accessibility of NGS technologies has allowed for the analysis of almost every molecular subtype, from DNA and RNA to proteins and metabolites.[227] However, major challenges can arise when attempting to combine or compare the results of these various 'omics analyses. For example, mRNA and protein expression levels are often not strongly correlated[184] making direct comparison and interpretation difficult.

One solution that aims to mitigate such issues is Similarity Network Fusion which integrates data by patient thereby having the number of predictors depend on the number of subjects rather than molecular features.[404] This method is well suited to integrate highly

heterogeneous 'omic data sets and has been used to integrate multi-omic profiles of patients with COPD compared to smokers and non-smokers with normal lung function.[234]

By combining knowledge from the fields of biology and computer science together, it is possible to store and analyze large amounts of 'omic data in novel ways. Integration of multi-omic profiles can provide insights into the vast connections between genes and environmental factors that impact risk of disease onset and progression.[184] It is also through addressing this biological complexity and through the close collaboration of basic and clinical scientists with bioinformaticians that precision medicine can become a reality.[13]

1.3 Thesis outline

The broad theme of the work presented in this thesis is the analysis of the effects of various external stimuli on the human airway using bioinformatic techniques. My research aims to characterize these effects in the hopes of better informing public health policy, increasing scientific knowledge, and improving patient outcomes through precision prevention, detection, and treatment. Throughout this thesis I explore my research questions using a combination of focused bioinformatic studies of original sequencing data sets as well as large-scale analyses of publicly available studies, demonstrating the value of multi-omic data analyses to respiratory medicine.

The thesis is structured as five chapters. This first introductory chapter provides biological and methodological context to aid interpretation of the following data chapters. The three data chapters are organized by inhaled particulate.

Chapter 2 presents analyses of tobacco smoke and viral mimic. This chapter addresses the effects of these stimuli in a directed manner by focusing on a particular gene family of interest: the ABC transporters, which are ubiquitous within the human airway and have been previously suggested to have involvement in lung function and health. This chapter is broken down further into two sub-chapters. Chapter 2.1 characterizes the effects of tobacco smoking and the onset of common lung diseases such as COPD and asthma on the expression of ABC transporters in the human airway. Chapter 2.2 focuses on a single ABC transporter (ABCF1) and how it may be involved in the viral-induced immune response of the airway epithelium. The goal of this chapter is to interrogate well-studied particulates (tobacco smoke and virus) in a new way by using bioinformatic approaches to characterize the previously unexplored effects of these inhaled stimuli.

Chapter 3 focuses on cannabis smoke exposure and how it relates to tobacco smoke exposure in the context of functional outcome, response to combination treatment, and

association with disease. This data chapter is divided into three sub-chapters which respectively focus on the comparison between cannabis smoke- and tobacco smoke-induced host responses, the efficacy of long-acting β-agonist/glucocorticoid (LABA/GC) intervention in the context of cannabis and tobacco smoke exposures, and the association between the endocannabinoid system and lung diseases such as COPD and asthma. This chapter aims to expand upon currently available research on the effects of cannabis smoke exposure in the human airway to increase collective knowledge on this increasingly prevalent stimulus.

Chapter 4 is related to SARS-CoV-2, the novel coronavirus responsible for the COVID-19 pandemic. This chapter aims to characterize the mechanism for coronavirus-host cell entry (Chapter 4.1) as well as the subsequent host immune response triggered upon infiltration (Chapter 4.2). The goal of this chapter is to contribute to the global efforts in understanding and combating this novel virus while simultaneously preparing for the emergence of future disease pandemic-causing pathogens.

Finally, Chapter 5 features a discussion of general and specific information gained from preceding chapters, including technical and methodological considerations that affect results and their interpretation, and concludes with a short prospective for the field.

Chapter 2

Tobacco smoke exposure and viral mimic

Of the many compounds to which the human respiratory system may be exposed, tobacco cigarette smoke and respiratory viruses are among the most prevalent and well-studied. While the causal links between exposure to these particulates and the onset of various lung diseases are accepted, there is still much that is not yet understood regarding the mechanisms governing how epithelial cells respond to these stimuli. With the use of bioinformatic tools and techniques, researchers are able to hone in on previously under-explored areas of even the most well-characterized airway exposures, leveraging large repositories of existing data cohorts to further advance our collective understanding of the effects tobacco smoke and common respiratory tract viral infections have on lung biology.

A relatively unexplored contributor to respiratory mucosa biology is the ATP-binding cassette (ABC) family of transporters.[81,90,120,132,181,197,357,393] ABC transporters are ubiquitous across all three domains of life – Archaea, Bacteria, and Eukarya,[395] with humans expressing 48 transporters divided into 7 classes based on structure and function (ABCA, ABCB, ABCC, ABCD, ABDE, ABCF, and ABCG).[323,396] The majority of ABC transporters couple ATP hydrolysis to the extracellular transport of substrates. ABC transporters have been demonstrated to transport cytokines, ions, and lipids, and ABC transporter expression in human airway epithelial cells may contribute to respiratory mucosal immunity by modulating the lung environment in response to environmental exposures including cigarette smoke, allergens, air pollution, bacteria, and viruses, although the confirmation of these functions and expression of these genes and proteins has not been extensively confirmed in the respiratory mucosa of humans.[323,396]

The precedent for the importance of ABC transporters in respiratory mucosa biology has been set by the identification of mutations in ABC transporter genes in patients with lung pathologies. *ABCC7*, also known as cystic fibrosis transmembrane conductance regulator (*CFTR*), is the causal gene contributing to the development of cystic fibrosis.[14,348] Over 2000 variants in CFTR have been reported, resulting in altered chloride and bicarbonate secretion, hydration of the airway surface lining fluid, and mucus function.[14,348] Equally convincing of the role of ABC transporters in lung health and disease is the link between genetic loss of function variants of *ABCA3* and fatal surfactant deficiency in newborns.[359] In addition, members of the Hirota research group have recently confirmed that *ABCC4* is expressed in airway epithelial cells,[90,132] is able to transport cAMP and uric acid,[132] and can modulate anti-inflammatory activities of long-acting β-agonist/glucocorticoid therapies[181] and potentiate CFTR in select patient populations.[151] These three mechanistic and clinically relevant examples represent only a fraction of the 48 transporters that remain to be more extensively studied in the respiratory mucosa for expression patterns and function in lung health and disease.

2.1 The impact of cigarette smoke exposure, COPD, or asthma status on ABC transporter gene expression in human airway epithelial cells

Material in this section has been published as part of Aguiar *et al.* (2019).[11] The published manuscript is available here:

> J. A. Aguiar, A. Tamminga, B. Lobb, R. D. Huff, J. Nguyen, Y. Kim, A. Dvorkin-Gheva, M. R. Stämpfli, A. C. Doxey, and J. A. Hirota. The impact of cigarette smoke exposure, COPD, or asthma status on ABC transporter gene expression in human airway epithelial cells. *Scientific Reports*, 9(1):153. 2019.[11] https://doi.org/10.1038/s41598-018-36248-9

2.1.1 Introduction

The tobacco epidemic constitutes a major global public health crisis. Approximately 8 million people around the world die from tobacco-related exposures annually with greater than 7 million of those deaths being caused by direct tobacco use.[2] Tobacco smoke exposure

is the leading cause of preventable death worldwide and has been well-characterized as causing various lung diseases such as lung cancer and chronic obstructive pulmonary disease (COPD) such as bronchitis and emphysema.[277] Tobacco smoke exposure, including secondhand exposure, is also a common trigger for asthma.[1] While the effects of tobacco smoke exposure are widely understood, computational methods can still be utilized to uncover further information that may aid in the prevention and treatment of tobacco-induced airway diseases.

As previously mentioned, confirmation of ABC transporter expression in the human airway is currently lacking. Therefore, I have initiated a characterization of all 48 ABC transporters in the respiratory mucosa by performing a bioinformatic analysis of 9 distinct gene-expression data sets of primary human airway epithelial cells isolated from bronchial brushings of well-phenotyped healthy subjects and individuals that smoke cigarettes, or have been diagnosed with COPD or asthma.[87,151,222,263,312,359,364,383,418] My hypothesis was that specific ABC transporter gene expression patterns would correlate with presence of a specific chronic respiratory disease and disease severity. To validate the bioinformatic analyses, I interrogated select results using an *in vitro* cell culture model system to provide a foundational platform for further interrogation of candidate ABC transporters identified that may be important in chronic respiratory disease pathology. *In situ* human gene expression data analysis reveals that ABC transporters are i) variably expressed in epithelial cells from different airway generations, ii) responsive to cigarette smoke exposure, iii) differentially expressed in individuals with COPD and asthma, and iv) are amenable to interrogation with *in vitro* cell culture systems. I conclude that continued research into the basic biology of ABC transporters in the respiratory mucosa with mechanistic approaches is required to truly define the importance of these molecules in lung health and disease.

2.1.2 Methods

Data set selection and quality control

Analysis was performed on previously deposited data sets obtained from the Gene Expression Omnibus (GEO). Data sets pertaining to smoking and COPD analysis include GSE994, GSE4498, GSE11784, GSE11906, and GSE37147. Data sets focusing on asthma and severity include GSE4302, GSE63142, GSE67472, and GSE76227.[87,151,222,263,312,359,364,383,418]

Since the data sets were generated from a variety of studies, microarray chip selection and data normalization methods vary across the data sets. GSEs 11906, 11784, 4498, and 994 used the MAS 5.0 normalization method without $\log_2$-transformation, whereas GSEs

76227, 4302, 67472, and 37147 used the Robust Multi-array Average (RMA) method with $\log_2$-transformation. GSE63142 used the cyclic LOESS normalization method. To avoid applying multiple compounding normalization methods, data sets were left as provided by the original studies and differences in methods noted, with specific mention of the normalization method denoted in each figure axis. Since probe set IDs and corresponding gene targets vary across microarray platforms, expression results are reported in the context of both ABC transporter name and probe set ID. All data set demographics are available in Table 1.

ABC transporter transcript expression in human airway epithelial cells

Differential expression patterns for all 48 ABC transporter members were examined across the trachea, large airways (generation 2nd-3rd), and small airways (generation 10th-12th) using GSE11906 (Affymetrix Human Genome U133 Plus 2 microarray platform).[312] A bar plot showing average expression of ABC transporters in healthy, non-smoker lung tissue was generated using the ggplot2 package in R (version 3.4.3). In GSE11906, independent samples of epithelial cells were isolated from healthy individuals with no smoking history and normal lung function and included 17 trachea, 21 large airway, and 35 small airway samples.[359]

Impact of cigarette smoke exposure on ABC transporter expression

Three data sets (GSE11906, GSE11784, and GSE4498), which were all generated from the Affymetrix Human Genome U133 Plus 2 microarray platform, were used to assess the impact of cigarette smoke exposure on ABC transporter gene expression.[151,312,383] All three data sets are comprised of small airway (10th-12th generation) epithelial cell transcript expression patterns in healthy subjects and those with a history of smoking without a diagnosis of COPD. In GSE11906, 54 independent samples of epithelial cells were isolated from individuals with $>$ 25 pack years smoking history with no reported COPD.[312] In GSE11784, 72 independent samples of epithelial cells were isolated from individuals with $>$ 25 pack years smoking history with no reported COPD.[383] In GSE4498, 10 independent samples of epithelial cells were isolated from individuals with $>$ 25 pack years smoking history with no reported COPD.[151] These three GSE data sets were independently curated to ensure that no samples were repeated across the analyses.

Impact of cigarette smoking cessation on ABC transporter expression

Two data sets (GSE37147 and GSE994) generated from two distinct microarray platforms (Affymetrix Human Gene 1 ST and Affymetrix Human Genome U133A, respectively), which analyzed epithelial cells from medium (6th-8th generation) and large (2nd generation) airways, respectively, were used to assess the impact of smoking cessation on ABC transporter gene expression.[359,364] In GSE37147, independent samples of epithelial cells were isolated from 69 current smokers and 82 former smokers with $>$ 47 pack years smoking history with no reported COPD.[364] In GSE37147, the average duration of smoking cessation was 11.11 years. In GSE994, independent samples of epithelial cells were isolated from 34 current smokers, 14 former smokers, and 23 never-smokers, with $>$ 22 pack years smoking history with no reported COPD.[359] In GSE994, the average duration of smoking cessation was 10.49 years.

Association between COPD status and changes in ABC transporter expression

Three data sets (GSE11906, GSE11784, and GSE37147) were used to assess association of COPD status with ABC transporter gene expression profile. GSE11906 and GSE11784 collected epithelial cells from small airways (10th-12th generation) while GSE37147 collected from medium airways (6th-8th generation), with the two different sample types also analyzed on different microarray platforms (Affymetrix Human Genome U133 Plus 2 and Affymetrix Human Gene 1 ST, respectively).

In GSE11906, 20 independent samples of epithelial cells were isolated from individuals with $>$ 38 pack years smoking history with reported COPD and were compared to 54 independent samples of epithelial cells isolated from individuals with $>$ 25 pack years smoking history with no reported COPD.[312] In GSE11784, 36 independent samples of epithelial cells were isolated from individuals with $>$ 34 pack years smoking history with reported COPD and were compared to 72 independent samples of epithelial cells isolated from individuals with $>$ 25 pack years smoking history with no reported COPD.[383] In GSE37147, 87 independent samples of epithelial cells were isolated from individuals with $>$ 51 pack years smoking history with reported COPD and were compared to 151 independent samples of epithelial cells isolated from individuals with $>$ 47 pack years smoking history with no reported COPD.[364]

ABC transporter expression patterns in airway epithelial cells from asthmatics

For asthma-focused analyses, two pairs of data sets were used: GSE4302 and GSE67472,[87,263,418] which allowed for healthy control comparison to individuals with asthma, and GSE63142 and GSE76227,[87,222,263] which allowed for associating expression profiles with asthma severity. All studies analyzed epithelial cells from medium (3rd-5th generation) airways, while distinct microarray platforms were used (GSE4302 and GSE67472 used Affymetrix Human Genome U133 Plus 2, GSE63142 used Agilent 014850 Whole Human Genome Microarray 4 x 44K, and GSE76227 used Affymetrix HT HG U133 plus PM).

In GSE4302, independent samples of epithelial cells were isolated from 28 healthy individuals and 42 asthmatics.[418] In GSE67472, independent samples of epithelial cells were isolated from 43 healthy individuals and 62 asthmatics.[87] In GSE63142, independent samples of epithelial cells were isolated from 26 healthy individuals, 59 mild asthmatics, 19 moderate asthmatics and 51 severe asthmatics.[263] In GSE76227, independent samples of epithelial cells were isolated from 26 healthy individuals, 59 mild asthmatics, 19 moderate asthmatics and 51 severe asthmatics.[222]

In vitro validation of candidate ABC transporter gene expression changes

In vitro validation performed by Ryan Huff at The University of British Columbia

A cigarette smoke extract conditioned media experiment with the Calu-3 airway epithelial cell line was performed as previously described.[180] Expression of *ABCA13* and *ABCC1* was validated by quantitative-PCR. Cell viability was determined using an LDH assay.

Data set compiling and statistical analyses

Curated data sets contained the data matrix and relevant categorical data such as smoking status, disease status or airway location from which samples were isolated. Lists of probe set IDs for each microarray platform used in the study were generated using only probes that corresponded to ABC transporter proteins. Lists were used to filter corresponding matrices to include data exclusively for probes that target ABC transporters across data sets.

Pairwise Mann-Whitney U tests were performed to assess the statistical significance of differential gene expression between conditions. Adjusted p values (q values) were derived using the Benjamini-Hochberg multiple hypothesis testing correction as implemented in R (version 3.4.3). This statistical method was chosen because *a priori* analysis of all data

set distributions showed that the data was non-Normal. Non-parametric methods such as Mann-Whitney U (which is the non-parametric equivalent of the unpaired Student's T-test) make no assumptions about distribution and are therefore preferred for the analysis of non-Normal data.

Box plots were generated for individual probe sets using the ggplot2 package in R (version 3.4.3). Box plots visualize the minimum value (bottom of vertical line), first to third quartile (box), median (horizontal line within box), and maximum value (top of vertical line) of a data distribution. To visualize the underlying raw data distributions, individual data values were plotted as points on top of each box plot with randomized horizontal skew.

For GSE994 which includes more than 2 independent categorical variables (current, former and never smokers) I used the Kruskal-Wallis test. The Kruskal-Wallis test is a non-parametric alternative to the one-way analysis of variance (ANOVA) test that does not assume a data distribution. The Tukey Honest Significant Difference (Tukey HSD) test was used to perform *post-hoc* pairwise comparisons with multiple testing corrections on probe set IDs found to be significantly differentially expressed across categories as determined by the Kruskal-Wallis analysis.

To summarize the performed statistical analyses, tables were generated for each data set that included Probe set ID, Probe set, the Benjamini-Hochberg corrected p value (q), and the $\log_2$ fold change. The top 10 probe set IDs were listed for each data set based on corrected p value, and multiple data sets used to analyze the same biological question were combined into final summary tables.

For the *in vitro* study of cigarette smoke extract exposure, a one-way ANOVA was performed with a Bonferroni correction for multiple comparisons with a significance determined at $p < 0.05$.

2.1.3 Results

ABC transporter transcript expression in human airway epithelial cells

To begin the characterization of the 48 known ABC transporters in airway epithelial cells, I examined expression patterns of each member across the trachea, large airways (generation 2nd-3rd), and small airways (generation 10th-12th) (Figure 2.1; Table 2) using GSE11906. A global assessment of all ABC transporters across airway generations revealed diversity in their expression, with the highest detected expression levels for *ABCA5*, *ABCA13*, and *ABCC5* (Figure 2.1) and a trend for increased expression in the trachea relative to the

small airways. Between the airway generations, differences were only observed between the small airways and the trachea ($q < 0.05$ - Table 2). The top three candidates ranked by $\log_2$fold-change in small airway relative to trachea were *ABCC1*, *ABCC4*, and *ABCB3* (all with $q < 0.05$). Collectively the data suggest that ABC transporters are expressed in human airway epithelial cells with airway generation specific expression patterns.

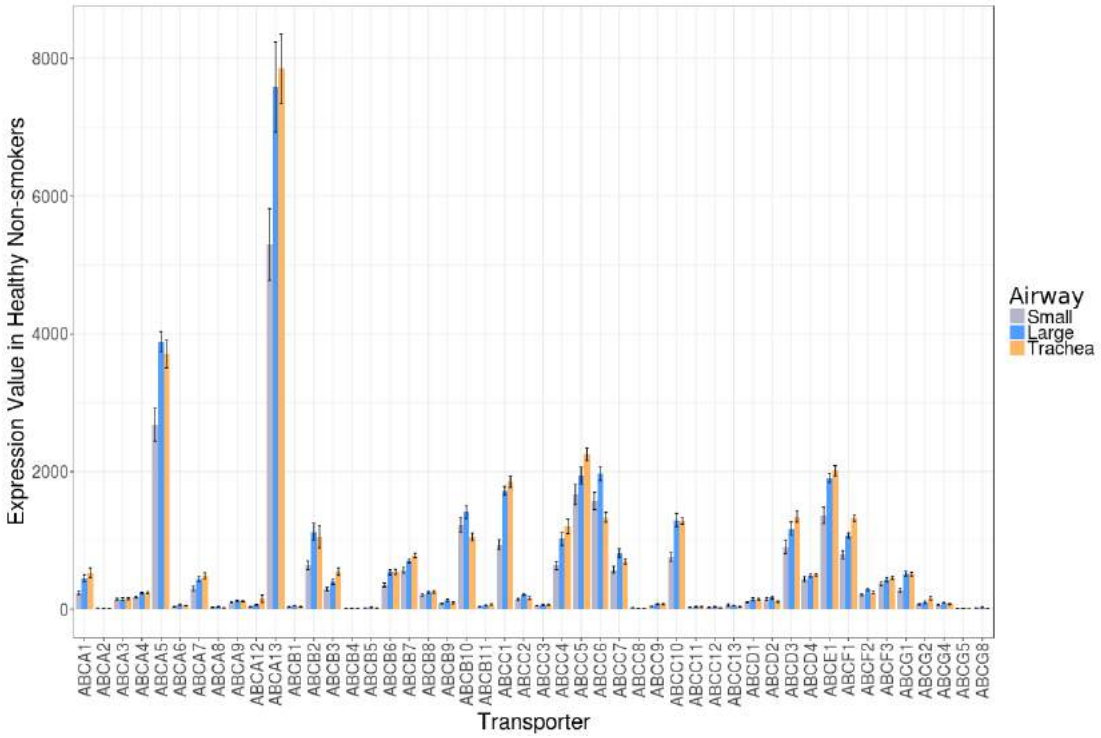


Figure 2.1: Gene expression levels for all 48 ABC transporter members were examined across the trachea, large airways (generation 2nd-3rd), and small airways (generation 10th-12th) using a single data set (GSE11906, Affymetrix Human Genome U133 Plus 2 microarray platform) comprised of independent epithelial cell samples isolated from healthy, non-smoking individuals. A representative probe set for each ABC transporter was chosen and average signal intensities (with standard error) for each probe set were calculated and visualized as a bar plot using ggplot2 in R (version 3.4.3). grey: small airways, blue: large airways, orange: trachea.

Impact of cigarette smoke exposure on ABC transporter expression

I next explored the impact of cigarette smoke exposure on ABC transporter expression profiles. For my analysis I interrogated three data sets of small airway (10th-12th generation) epithelial cells from healthy subjects and those with a history of smoking without a diagnosis of COPD (GSE11906, GSE11784, and GSE4498).[151,312,383]

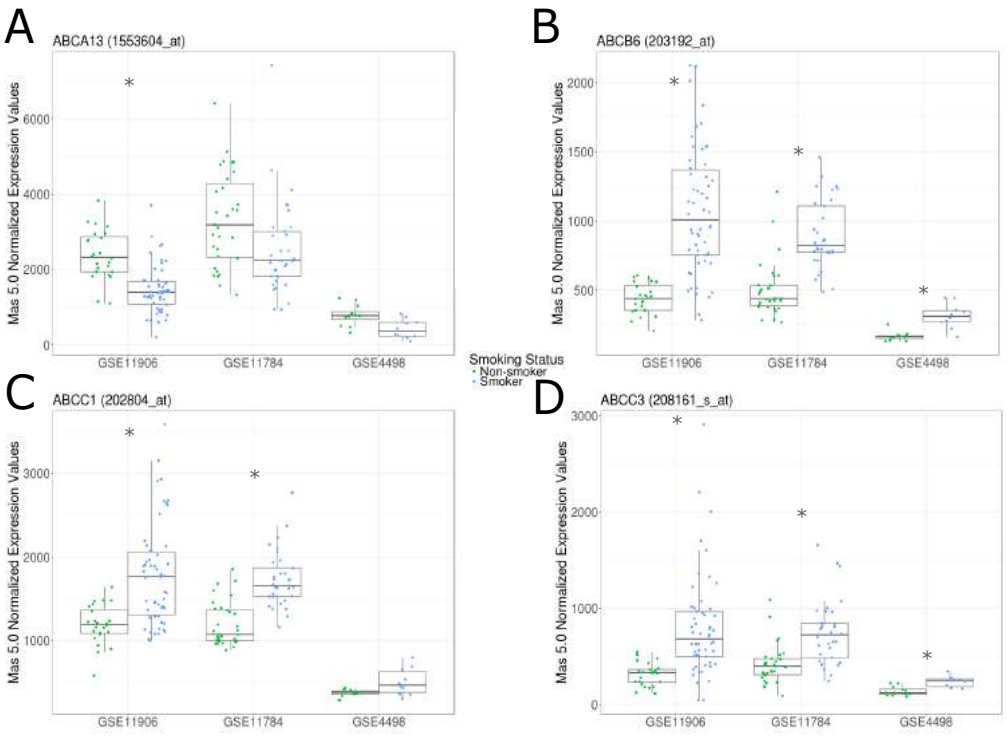


Figure 2.2: Gene expression levels of selected ABC transporters ((A) *ABCA13*, (B) *ABCB6*, (C) *ABCC1*, (D) *ABCC3*) between healthy non-smokers (green) and smokers with no diagnosis of COPD (blue). Significant expression differences are indicated by asterisks ($q < 0.05$ according to Mann-Whitney U test with Benjamini-Hochberg correction). All three data sets used for this analysis (GSE11906, GSE11784, and GSE4498) analyzed epithelial cells from small airways (10th-12th generation) and were generated from the Affymetrix Human Genome U133 Plus 2 microarray platform (using the MAS 5.0 normalization method). Data distributions are visualized as box plots (see Methods for additional details).

Cigarette smoke exposure was associated with elevated *ABCB6* and *ABCC3* expression levels, a feature conserved across all three independent data sets (Figure 2.2; Table 3). In two of three data sets, an elevation in *ABCC1* was also observed, with similar non-significant trends observed in the third data set. In one of three data sets, lower *ABCA13* expression was observed, with similar non-significant trends observed in the two other data sets.

Impact of cigarette smoking cessation on ABC transporter expression

Cigarette smoking is a modifiable risk factor for the development of COPD with cessation advocated by the Global Initiative for Obstructive Lung Disease.[398] I therefore determined whether smoking cessation could normalize differences observed in ABC transporter expression for *ABCA13*, *ABCB6*, *ABCC1*, and *ABCC3* by comparing former smokers, current smokers, and where possible, healthy never-smokers. For my analysis I examined two data sets generated from two distinct microarray platforms, GSE37147 and GSE994, which analyzed epithelial cells from medium (6th-8th generation) and large (2nd generation) airways, respectively.

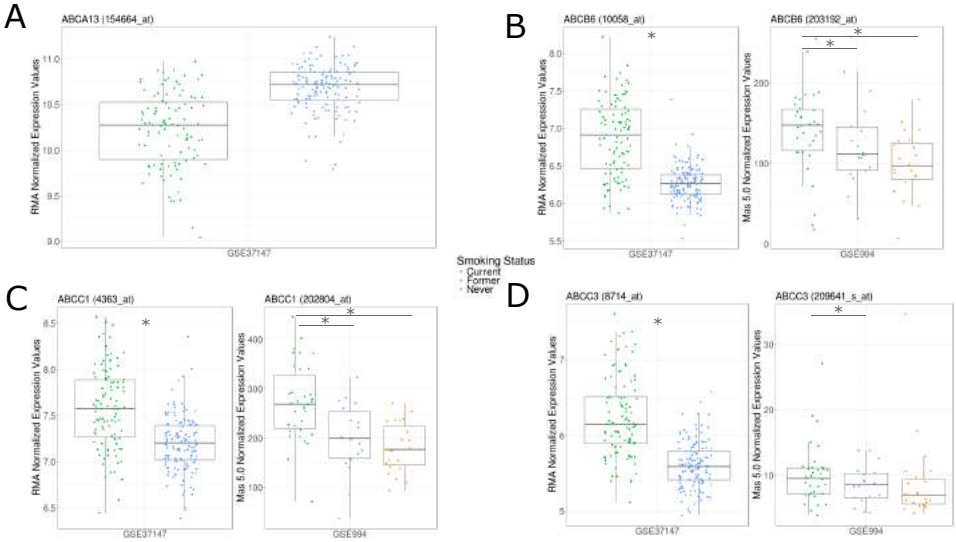


Figure 2.3: Differential expression of (A) *ABCA13* between current smokers (green) and former smokers (blue) from GSE37147 (*: $q < 0.05$, Mann-Whitney U pairwise statistical analysis with Benjamini-Hochberg multiple testing correction). GSE37147 was selected as it was the only data set including a probe matching *ABCA13*. Differential expression of (B) *ABCB6*, (C) *ABCC1*, and (D) *ABCC3* between current smokers (green), former smokers (blue), and, where possible, never-smokers (orange) from GSE37147 and GSE994 (*: $q < 0.05$, Kruskal-Wallis followed by the Tukey Honest Significance Difference (Tukey HSD) test). The data sets used in this analysis (GSE37147 and GSE994) were generated from two distinct microarray platforms (Affymetrix Human Gene 1 ST and Affymetrix Human Genome U133A, respectively), which analyzed epithelial cells from medium (6th-8th generation) and large (2nd generation) airways, respectively. GSE37147 used the logged (base 2) RMA normalization method and GSE994 used the MAS 5.0 normalization method. Data distributions are visualized as box plots (see Methods for additional details).

Smoking cessation was associated with lower expression of *ABCB6*, *ABCC1*, and *ABCC3* in both data sets (Figure 2.3; Table 4). A trend for elevated *ABCA13* expression ($q = 0.06$) was observed in association with smoking cessation in the medium airways with no primer sets available in the large airway data set. In GSE994, an additional group of never-smokers reveals that levels of *ABCB6* and *ABCC1* elevated with smoking do not normalize to never-smoker levels even with prolonged smoking cessation (average 10.7 years).[359]

Association between COPD status and changes in ABC transporter expression

I next analyzed the expression profiles of ABC transporters in relation to confirmed COPD status in three data sets (GSE11906, GSE11784, and GSE37147). GSE11906 and GSE11784 collected epithelial cells from small airways (10th-12th generation) while GSE37147 collected from medium airways (6th-8th generation).

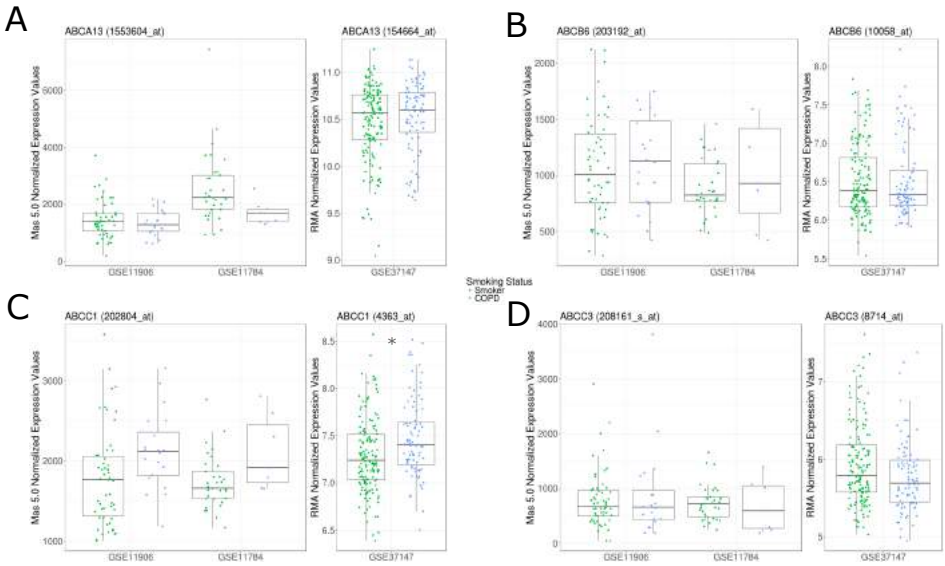


Figure 2.4: Gene expression levels of selected ABC transporters ((A) *ABCA13*, (B) *ABCB6*, (C) *ABCC1*, (D) *ABCC3*) between smokers with COPD (blue) and smokers with no diagnosis of COPD (green). Significant ABC transporter expression differences are indicated by asterisks ($q < 0.05$ according to Mann-Whitney U test with Benjamini-Hochberg correction). Three data sets (GSE11906, GSE11784, and GSE37147) were used for this assessment. GSE11906 and GSE11784 collected epithelial cells from small airways (10th-12th generation) while GSE37147 collected from medium airways (6th-8th generation), with the two different sample types also analyzed on different microarray platforms (Affymetrix Human Genome U133 Plus 2 and Affymetrix Human Gene 1 ST, respectively). GSEs 11906 and 11784 used the MAS 5.0 normalization method, and GSE37147 used the logged (base 2) RMA normalization method. Data distributions are visualized as box plots (see Methods for additional details).

No ABC transporter showed differential expression across all three independent data

sets (Figure 2.4; Table 5). Of candidates modified by cigarette smoking history, *ABCC1* was ranked in the top 10 differentially expressed probe sets in each data set containing smokers with and without COPD, with a trend observed for increased expression with COPD status that was significant in the medium airways (GSE37147) (Figure 2.4). *ABCA13*, *ABCB6*, and *ABCC3* expression profiles did not show any association with COPD status.

ABC transporter expression patterns in airway epithelial cells from asthmatics

For the asthma-focused analysis I interrogated two pairs of data sets: GSE4302 and GSE67472, which allowed for healthy control comparison to individuals with asthma and GSE63142 and GSE76227, which allowed for associating expression profiles with asthma severity.[87,222,263] All studies analyzed epithelial cells from medium (2nd-5th generation) airways.

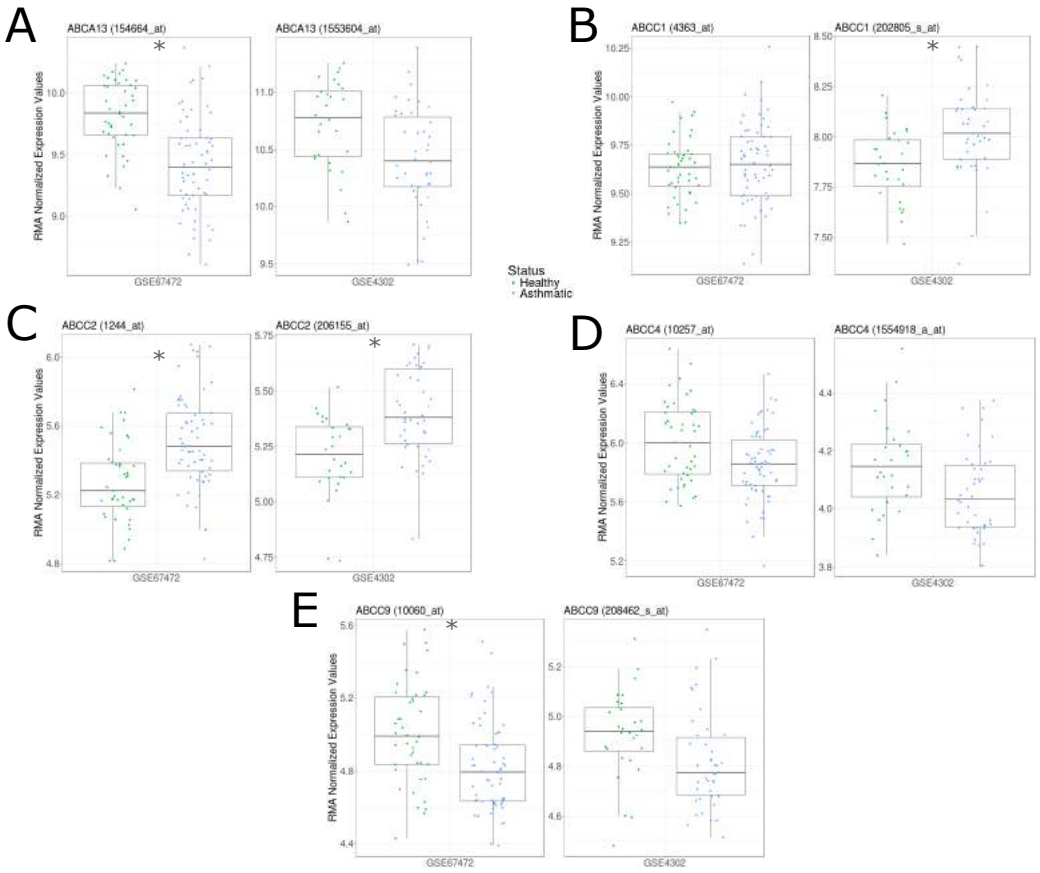


Figure 2.5: Gene expression levels of selected ABC transporters ((A) *ABCA13*, (B) *ABCC1*, (C) *ABCC2*, (D) *ABCC4*, (E) *ABCC9*) between healthy controls (green) and asthmatics (blue). Significant ABC transporter expression differences are indicated by asterisks ($q < 0.05$ according to Mann-Whitney U test with Benjamini-Hochberg correction). Both data sets (GSE4302 and GSE67472) analyzed epithelial cells from medium (3rd-5th generation) airways and were generated from the Affymetrix Human Genome U133 Plus 2 microarray platform (using the logged (base 2) RMA normalization method). Data distributions are visualized as box plots (see Methods for additional details).

It was observed that asthma status is associated with an increase in *ABCC2* in both data sets (Figure 2.5; Table 6). *ABCA13* and *ABCC9* were decreased in one data set (GSE67472) with a similar trend in the second data set ($q < 0.1$ – GSE4302). *ABCC1* expression was increased in one data set (GSE4302) with no trend observed in the second

data set. *ABCC4* expression levels showed trends towards reduced expression in samples from asthmatics in both data sets that were not significant.

In the data sets that stratified asthmatics based on disease severity (GSE63142 and GSE76227), trends for reduced *ABCC4* and *ABCA13* gene expression were associated with increased severity of disease (Figure 2.6; Table 7). In contrast, *ABCC1*, *ABCC2*, and *ABCC9* failed to trend with disease severity (data not shown).

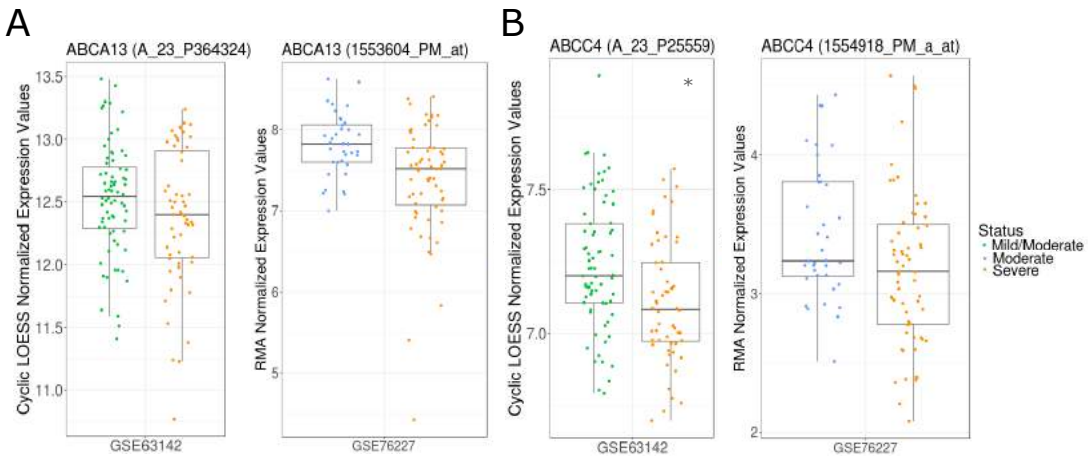


Figure 2.6: Gene expression levels of (A) *ABCA13* and (B) *ABCC4* between mild/moderate (green), moderate (blue), and severe (orange) asthmatics visualized as box plots (see Methods for additional details). Significant ABC transporter expression differences are indicated by asterisks ($q < 0.05$ according to Mann-Whitney U test with Benjamini-Hochberg correction). Both data sets used for this analysis (GSE63142 and GSE76227) analyzed epithelial cells from medium (3rd-5th generation) airways; however, they were generated using different microarray platforms (Agilent 014850 Whole Human Genome Microarray 4x44K G4112F and Affymetrix HT HG U133 plus PM, respectively). GSE63142 uses the Cyclic LOESS normalization method and GSE76227 uses the RMA normalization method (both methods use $\log_2$ scales).

In vitro validation of candidate ABC transporter gene expression changes

A priori it was determined that a complete *in vivo* COPD and asthma phenotype was unable to be recapitulated *in vitro*, but it was possible to recapitulate the impact of cigarette smoke exposure.[180] A cigarette smoke extract conditioned media experiment was therefore performed with the Calu-3 airway epithelial cell line grown under submerged

monolayer conditions to validate an *in vitro* model system for interrogating candidate ABC transporters of interest. Cigarette smoke extract conditioned media exposure (10% and 20%) for 24 hours recapitulates the decrease in *ABCA13* gene expression and increase in *ABCC1* gene expression that is observed in human bronchial brushing samples (Figure 2.7 A, B) from individuals that have smoked cigarettes (Figure 2.2 A, C), with no changes in cell viability Figure 2.7 C).

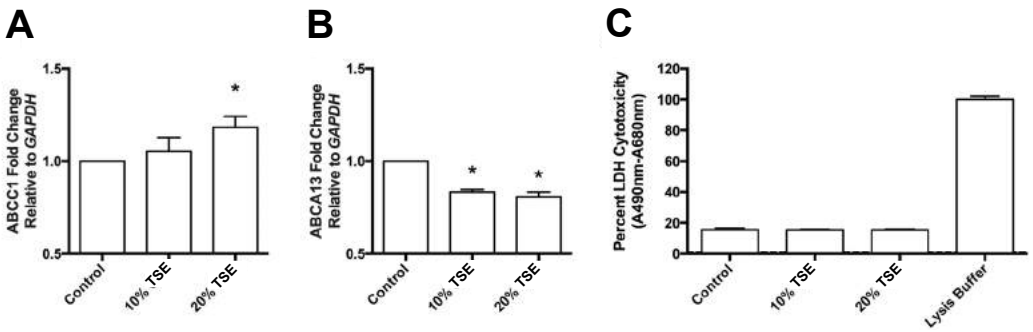


Figure 2.7: Expression levels of the *ABCA13* and *ABCC1* genes in Calu-3 cells grown under submerged monolayer conditions were assessed after exposure of cells to 10% and 20% tobacco cigarette smoke extract (TSE) conditioned media for 24 hours. Expression is relative to GAPDH levels and also compared to a negative control (unexposed Calu-3 cells) (A) *ABCC1* expression increases under these conditions, whereas (B) *ABCA13* expression decreases as % TSE increased with no observed impact on cell viability (C). One-way ANOVA was performed with a Bonferroni correction for multiple comparisons; $n = 3$; *: $p < 0.05$ relative to control.

2.1.4 Discussion

The contribution of ABC transporters to respiratory mucosal immunology remains to be defined. To address this knowledge gap, I performed a bioinformatic analysis of 9 distinct gene expression data sets of primary human airway epithelial cells to test the hypothesis that specific ABC transporter gene expression patterns would correlate with presence of a specific chronic respiratory disease and disease severity. The *in situ* human gene expression data demonstrate that ABC transporters are i) variably expressed in epithelial cells from different airway generations (top three expression levels - *ABCA5*, *ABCA13*, and *ABCC5*), ii) regulated by cigarette smoke exposure (*ABCA13*, *ABCB6*, *ABCC1*, and *ABCC3*), and iii) differentially expressed in individuals with COPD and asthma (*ABCA13*, *ABCC1*,

ABCC2, *ABCC9*). Lastly, this study demonstrates that an *in vitro* cell culture system of cigarette smoke exposure is amenable to investigate the consequences of differential gene expression patterns of candidate ABC transporters (*ABCA13* and *ABCC1*). This work sets a foundation for further mechanistic research into the basic biology of ABC transporters in the respiratory mucosa and suggests a potential contribution to chronic respiratory diseases.

My analysis on ABC transporter expression profiles across airway generations was performed on samples isolated from healthy subjects at the level of the trachea, large airways (generation 2nd-3rd), and small airways (generation 10th-12th). The data demonstrate that the expression levels of select ABC transporters are heterogeneous throughout the airway tree (e.g. *ABCC1*, *ABCC4*, and *ABCB3*) while others are homogenous (e.g. *ABCA13*, *ABCC2*, and *ABCC3*). The relative paucity of literature mechanistically describing the function of each ABC transporter in human airway epithelial cells precludes us from attributing any differential expression pattern to airway generation specific functional responses of the epithelium. Regardless, the data suggests that ABC transporter expression is differentially expressed along the airway tree, which may be important for chronic airway diseases that lead to pathology of distinct airway regions (e.g. COPD as a small airways disease[257]).

Environmental exposures including cigarette smoke, air pollution, allergens, viruses, and bacteria can induce immune mediator release in the airway epithelium.[164,288] ABC transporters may be responsible for immune mediator release, including uric acid, leukotrienes, prostaglandins, and glutathione conjugates,[81,120] that may be important for responding to environmental stimuli. To investigate possible ABC transporter-environment relationships, I explored the impact of cigarette smoke exposure and smoking cessation on *in situ* gene expression in epithelial cells. The analyses of the impact of smoking used three data sets generated on the same microarray platform (GSE11906, GSE11784, and GSE4498),[151,312,383] while the smoking cessation analysis used two additional data sets generated on two different microarray platforms (GSE37147 and GSE994).[359,364] Conserved elevations in *ABCB6*, *ABCC1*, and *ABCC3* gene expression in response to cigarette smoke exposure were observed, accompanied by trends for decreased gene expression with smoking cessation. *ABCB6* is a broad-spectrum porphyrin transporter expressed in mitochondria and plasma membranes involved in heme biosynthesis and protective responses to reactive oxygen species.[219,290] *ABCB6* over-expression increases catalase gene expression and stability[250] and is able to inhibit arsenic cytotoxicity,[80] a carcinogenic component found in cigarette smoke.[286] Despite the potential protective consequences of elevated *ABCB6* expression that was observed in cigarette smokers, it is important to note that negative consequences may also arise as elevated ABCB6 levels may lead to resistance to chemotherapeutic agents.[60] Therefore, the observed up-regulation in *ABCB6* is likely an important initial protective response to

cigarette smoke exposure that may have untoward consequences of reducing responsiveness to chemotherapeutic agents for management of lung cancers. *ABCC1* and *ABCC3* also provide protective effects by transporting glutathione-conjugated anions from the cytosol to the extracellular compartment to prevent cellular accumulation of toxic metabolites.[231,433] Similar to *ABCB6*, *ABCC1*, and *ABCC3* are also capable of effluxing chemotherapeutics of broad function including anti-cancer functions. Collectively it is conceivable that *ABCB6*, *ABCC1*, and *ABCC3* gene expression levels are regulated by a feedback mechanism linked to oxidative stress and toxin exposure induced by cigarette smoke exposure to help facilitate anti-oxidant activities within airway epithelial cells.

Importantly, it was observed that as cigarette smokers progress to develop COPD, *ABCC1* gene expression is elevated compared to individuals that do not have COPD. Linked to cigarette smoking, elevated *ABCC1* gene expression has been observed in both non-small cell carcinoma and small-cell carcinoma lung cancers.[67] The contribution of *ABCC1* biology in response to cigarette smoke exposure and to the associated development of COPD and lung carcinoma is intriguing. An *in vitro* model of cigarette smoke extract conditioned media induction of *ABCC1* gene expression may be valuable in further interrogating this biology and the functional consequences in the context of COPD and lung carcinomas.

In contrast to elevated gene expression levels in response to cigarette smoke exposure and COPD status, I observed a conserved decrease in the expression of *ABCA13*. I also observed a decrease in *ABCA13* gene expression in asthmatics relative to healthy controls, suggesting this may be a non-specific response to an inflammatory lung environment. *ABCA13* is an intriguing candidate as the expression levels are very high relative to other ABC family members, and there are no reports of endogenous or exogenous substrates, expression patterns in epithelium, or associations with respiratory mucosal immunology. Reports show associations between *ABCA13* expression and cancers,[24,125,166] while *ABCA13* variants are associated with mental health abnormalities,[108,213] although no mechanisms have been defined for either of these observations. Other ABCA family members are involved in lipid transport and dysregulation of their pathways can result in lung inflammation.[78] To gain insight into *ABCA13* biology, one can examine other ABCA family members: *Abca1*-KO mice have impaired lipid transport resulting in reduced serum cholesterol and HDL, elevated intracellular lipid contents, and abnormal lung structure.[40] In contrast, mice over-expressing the human *ABCA1* gene showed reduced inflammation and pathology in a model of allergic lung inflammation.[96] In humans and mice, dysfunction of *ABCA3* results in surfactant deficiencies and fatal respiratory failure.[47,322,348] Based on function of these two ABCA family members in the lung, I hypothesize that *ABCA13* is important in lipid handling in airway epithelial cells, and when decreased in expression, this could result in inflammation, altered surfactant production, and impaired innate immune functions. The

relatively high expression levels of *ABCA13* gene throughout the airway generations and the known function of ABCA family members in lung biology warrants a further exploration of this candidate.

The observation that cigarette smoke exposure, and to a small extent COPD status, was associated with changes in ABC transporter gene expression warranted further exploration to determine if these observations were non-specific responses to inflammatory airway environments. Although changes in *ABCA13* gene expression associate with both cigarette smoking and asthma status, unique changes in ABC transporter expression profiles are observed in samples isolated from asthmatics. Elevated *ABCC2* expression in asthmatics relative to healthy subjects was observed while reductions in *ABCC9* were observed. *ABCC2* is similar in function to *ABCC1* with the shared ability to transport glutathione-conjugated xenobiotics and control intracellular oxidative stress.[207,224,380] Elevations in *ABCC2* that are observed in airway epithelial cells in asthmatics may aid in control of oxidative environments in the asthmatic lung and epithelial cells,[79,190,321] but could also lead to extracellular transport of pharmacological agents used to control asthma due to the broad specificity of this transporter. Similar to *ABCC1* in smokers, changes in *ABCC2* in asthmatics may be both beneficial (management of oxidative stress) and detrimental (reduction in intracellular bioavailability of pharmacological agents). In contrast to elevations in *ABCC2*, I observed reductions in *ABCC9* in airway epithelial cells from asthmatics. *ABCC9* is also known as the sulfonylurea receptor 2 (SUR2) protein and functions as an ATP sensitive K + channel that helps coordinate calcium levels in skeletal and cardiac muscle.[116,198] Reduced *ABCC9* levels could conceivably dysregulate K + and calcium concentrations within airway epithelial cells, which in turn could impact K + regulated mechanisms of epithelial cell migration, proliferation, and tissue repair.[129] Irrespective of the proposed functions of these ABC transporters that are differentially expressed in epithelial cells from asthmatics, it can be concluded that the different patterns observed in cells from asthmatics, smokers, or individuals with COPD suggests that regulation and function of ABC transporters is specific to a given chronic inflammatory lung disease or environmental insult.

This bioinformatic study has some limitations that need to be recognized due to the use of publicly available data sets in the NCBI GEO. The deposited data sets were developed from different microarray platforms, probe sets, and epithelial cells isolated from different airway generations. Where the information was available, I have attempted to clearly disclose this in the methods and correspondingly the results sections have axes of figures labeled based on the pre-processing analysis performed by the original contributors of the data sets (including different normalization methods). Irrespective of these limitations and the inability to control for differences in microarray platform and source of epithelial

cells, the observed common trends across independent data sets are emphasized, suggesting validity of my approach and the available data sets.

In closing, I have initiated a characterization of the 48 known ABC transporters in human airway epithelial cells in the context of chronic airway diseases. Using 9 distinct gene-expression data sets of primary human airway epithelial cells, I tested the hypothesis that specific ABC transporter gene expression patterns would correlate with presence of a specific chronic respiratory disease and disease severity. It is clear that ABC transporters are i) variably expressed in epithelial cells from different airway generations, ii) regulated by cigarette smoke exposure, and iii) differentially expressed in individuals with COPD and asthma, I further demonstrate that an *in vitro* cell culture system is amenable to investigate the consequences of differential expression patterns of candidate ABC transporters, creating a foundation for further mechanistic research into the basic biology of ABC transporters in lung health and disease.

2.2 ABCF1 regulates dsDNA-induced immune responses in human airway epithelial cells

Material in this section has been published as part of Cao *et al.* (2019).[71] The published manuscript is available here:

> Q. T. Cao, J. A. Aguiar, B. J.-M. Tremblay, N. Abbas, N. Tiessen, S. Revill, N. Makhdami, A. Ayoub, G. Cox, K. Ask, A. C. Doxey, and J. A. Hirota. ABCF1 Regulates dsDNA-induced Immune Responses in Human Airway Epithelial Cells. *Frontiers in cellular and infection microbiology*, 10:487. 2020.[71] https://doi.org/10.3389/fcimb.2020.00487

2.2.1 Introduction

Respiratory tract viral infections are collectively responsible for more than 2.5 million deaths per year globally and represent an economic burden on health care systems for all demographics.[130] In individuals with underlying chronic airway disease, respiratory tract viral infections increase frequency and severity of disease exacerbations, hospitalizations, and contribute to morbidity and mortality.[191,239,324,352,399] Understanding the mechanisms governing respiratory tract viral infections and host defence is essential for the future development of treatments aimed at minimizing the morbidity and mortality of these

pathogens. As demonstrated in Chapter 2.1, bioinformatic analysis of previously understudied gene families such as ABC transporters can be an important tool in achieving these goals.

The ABCF family members are unique in their structure and function as they lack transmembrane regions and therefore lack capacity for transport of substrates.[319,395] Of the ABCF family members, ABCF1 is most extensively characterized in eukaryotes, with functions ranging from initiation of mRNA translation, immune modulation, and nucleic acid sensing.[25,228,294,295,319,388] The diverse functions attributed to ABCF1 are physiologically important, as demonstrated by the embryonic lethality of homozygous deletion of ABCF1 in mice.[413] To date, nucleic acid sensing by ABCF1 has been defined using the dsDNA immunostimulatory DNA (ISD) sequence derived from *Listeria monocytogenes*[367] and a dsDNA HIV sequence, with both nucleic acid motifs inducing CXCL10, interferon-β expression, and downstream type I interferon responses in mouse embryonic fibroblasts.[228] Complementary to dsDNA sensing, immune modulation mediated by ubiquitin-conjugating activities of ABCF1 have been defined in the context of macrophage polarization and immune responses linked to interferon-β production and tolerance important in mouse models of sepsis.[25] In the context of studies using human lung samples, *ABCF1* gene expression has been identified in the human airway epithelium,[11] although confirmation of protein and function remains to be determined. The clear *in vivo* demonstration of ABCF1 functions in immune responses in mouse models and the presence of detectable *ABCF1* gene expression the human airways warrants a deeper interrogation into the expression and function of this molecule in human health and disease.

Defining defence mechanisms in airway epithelial cells has important consequences in both lung health and disease, with the potential for interventions that could reduce viral-induced pathologies and exacerbations of chronic respiratory diseases.[191,239,324,352,399] Therefore, the hypothesis that ABCF1 functions as a dsDNA nucleic acid sensor in human airway epithelial cells important in regulating antiviral responses was tested, using archived human lung samples and human airway epithelial cells. Expression and localization experiments were performed using *in situ* hybridization and immunohistochemistry in human lung tissue from healthy subjects, while confirmatory transcript and protein expression was performed in human airway epithelial cells. Functional experiments were performed with siRNA methods as no selective small molecule inhibitors to ABCF1 have been validated to date. Complementary transcriptomic analyses were performed to explore the potential contributions of ABCF1 beyond dsDNA virus sensing. These results confirm expression of ABCF1 in human airway epithelial cells with a role for mediating CXCL10 production in response to dsDNA viral mimic challenge. Although ABCF1 knockdown was associated with an attenuation of select genes involved in the antiviral responses, Gene Ontology analyses

revealed a greater interaction of ABCF1 with TLR signaling suggesting a multi-factorial role for ABCF1 in innate immunity in human airway epithelial cells.

2.2.2 Methods

Human ethics

All studies using primary human lung material were approved by Hamilton integrated Research Ethics Board (HiREB – 5305-T and 5099-T).

Reagents

In situ hybridization was performed using a custom RNAscope™ probe for *ABCF1* (construct targeting 1713-2726 of NM_001025091.1) generated by Advanced Cell Diagnostics (ACD, Newark, California). Negative and positive control probes for quality control of RNA signal in analyzed human tissues were purchased from ACD (data not shown). Protein cell lysates were collected by lysing and scraping cells with RIPA Lysis buffer (VWR, Mississauga, Ontario) mixed with protease inhibitor cocktail (Sigma-Aldrich, Oakville, Ontario). Immunoblots were conducted using Mini-Protean TGX stain-free gels and Transfer-Blot Turbo RTA Transfer Kit reagents (Bio-Rad, Mississauga, Ontario). ABCF1 protein was probed with primary anti-ABCF1 antibody (HPA017578, Sigma-Aldrich, Oakville, Ontario) at 1:100 in 3% Casein in 1X Tris Buffered Saline with TWEEN ®20 (Sigma-Aldrich, Oakville, Ontario, and Anti-rabbit HRP-linked Antibody (7074S - Cell Signaling Technology, Danvers, MA) at 1:2000. Immunohistochemistry was performed using the same anti-ABCF1 antibody as immunoblotting. ABCF1 and scramble siRNA SMARTpool siGENOME transfection reagents were purchased from Dharmacon (M-008263-01 and D-001206-14, Lafayette, Colorado). Cell viability was estimated with the Pierce LDH Cytotoxicity Assay kit (ThermoFisher Scientific, Mississauga, Ontario). RNA samples were lysed with Buffer RLT and purified with RNeasy Mini Kit columns (Qiagen, Toronto, Ontario). The ligands ISD, ISD control, VACV-70, VACV-70 control, and Poly:IC were all complexed with LyoVec transfection reagent and purchased from Invivogen (San Diego, California). Human CXCL10 was quantified using a commercial ELISA with ancillary reagent kit (R&D Systems, Oakville, Ontario). The protocol for quantifying CXCL10 was modified with the use of a loading plate for the samples.

Cell culture

Cell culture performed by Quynh Cao at McMaster University

All experiments were performed in submerged monolayer cell culture. An immortalized human airway epithelial cell line (HBEC-6KT) over expressing human telomerase reverse transcriptase (hTERT) and cyclin-dependent kinase 4 (Cdk4) was used as previously described.[162,165,180,181,313] HBEC-6KT were grown in keratinocyte serum free medium (ThermoFisher Scientific, Mississauga, Ontario) supplemented with 0.8 ng/mL epithelial growth factor, 50 µg/mL bovine pituitary extract and 1X penicillin/streptomycin. Calu-3 cells (ATCC HTB-55) were grown in Eagle's Minimum Essential Media supplemented with 10% fetal bovine serum (Wisent, Saint-Jean-Baptiste, QC), 1mM HEPES, and 1X penicillin/streptomycin (Sigma-Aldrich, Oakville, Ontario). Primary human bronchial epithelial cells derived from healthy patient bronchial brushings were grown in PneumaCult ExPlus Medium supplemented with 96 µg/mL hydrocortisone (StemCell Technologies, Vancouver, BC) and 1X antimicrobial-antimycotics (ThermoFisher Scientific, Mississauga, Ontario). All cells were grown at 37°C at 5% CO2. Experiments with primary cells were performed between passages 1 and 4, and experiments with HBEC-6KT and Calu-3 cells were performed within 5 passages.

In vitro experiments

In vitro experiments performed by Quynh Cao at McMaster University

All *in vitro* knockdown experiments in HBEC-6KT were done using siRNA transfected with DharmaFECT Transfection Reagent according to the manufacturer's instructions. Cells were transfected with siABCF1 or siCTRL for 24 hours. After knockdown, cells were transfected with an immunostimulatory ligand for 24 hours followed by outcome measurements of cell viability (LDH and cell morphology), function (CXCL10 secretion), protein expression (immunoblot), or gene transcription (microarray). For TNF-α stimulation experiments, 10 ng/mL was incubated for 24 hours followed by protein collection for immunoblots. For ISD and VACV-70 stimulation experiments, a concentration- response study was performed using 0.316-3.16 µg/mL (ISD) or 0.1-3.16 µg/mL (VACV-70) followed by incubation for 24 hours. For Poly I:C stimulation experiments, 1 µg/mL was incubated for 24 hours.

Cytokine analysis

Cell supernatants were collected following *in vitro* experiments and sent to Eve Technologies for a Human Cytokine Array/Chemokine Array 48-Plex (Eve Technologies, Calgary, Alberta). Eve Technologies uses the Bio-Plex ® 200 to detect 49 different cytokines, chemokines and growth factors: sCD40L, EGF, Eotaxin, FGF-2, Flt-3 ligand, Fractalkine, G-CSF, GM-CSF, GROα, IFNα2, IFNγ, IL-1α, IL-1β, IL-1ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12 (p40), IL-12 (p70), IL-13, IL-15, IL-17A, IL-17E/IL-25, IL-17F, IL-18, IL-22, IL-27, IP-10 (CXCL10), MCP-1, MCP-3, M-CSF, MDC (CCL22), MIG, MIP-1α, MIP-1β, PDGF-AA, PDGF-AB/BB, RANTES (CCL5), TGF-α, TNF-α, TNF-β, VEGF-A.

Histology, digital slide scanning and microscopy

Histology, digital slide scanning and microscopy performed by Quynh Cao at McMaster University

In situ hybridization and immunohistochemistry was performed using a Leica Bond Rx autostainer with instrument and application specific reagent kits (Richmond Hill, Ontario). The human lung tissues selected for analysis were formalin fixed, paraffin embedded, lung samples from archived hospital clinical samples, determined to be free of defined lung pathology. Following selection, four micron thick sections were stained with using RNAscope^{TM} probes (*in situ* hybridization) or antibody (immunohistochemistry) following directions supplied with the Leica Bond reagent kits. For IHC, heat-induced antigen retrieval in citrate buffer was performed at pH 6 with primary antibody diluted at 1:50. Slides underwent digital slide scanning using an Olympus VS120-L100 Virtual Slide System at 40X magnification with VS-ASW-L100 V2.9 software and a VC50 colour camera (Richmond Hill, Ontario). Image acquisition and formatting was performed using Halo Software (Indica Labs, Albuquerque, NM).

Gene Expression Omnibus (GEO) data set mining

Gene expression patterns of *ABCF1* in human airway epithelial cells were determined relative to markers for immune cells (*CD34*), ABC transporters of known function in airway epithelial cells (*ABCC4*, *ABCC7*), and junctions (*CDH1*) in a data set containing samples from trachea, large airways (generation 2nd-3rd), and small airways (generation 10th-12th) from healthy subjects (GSE11906, Affymetrix Human Genome U133 Plus 2 microarray

platform).[312] The following probe sets were used to extract gene expression data: *ABCF1* (200045_at), *ABCC4* (203196_at), *ABCC7* (*CFTR*; 205043_at), *CDH1* (201131_s_at), and *CD34* (209543_s_at). In cases where more than one probe corresponded to a given gene, the following hierarchy was used to select an individual probe for further use: perfect, unique matches (probes ending in _at or _a_at) were preferred over mismatch or non-unique probes (ending in _s_at or _x_at). GSE11906 included 17 trachea (age – 42 +/- 7), 21 large airway (age – 42 +/- 9), and 35 small airway samples.

Processing of raw microarray data

Raw intensity values from a microarray experiment using the Affymetrix Clariom S Human chip-type were imported into the R statistical language environment (version 3.6.1; R Core Team, 2019). Probe definition files were obtained from the Brainarray database (version 24[97]). The Single Channel Array Normalization (SCAN) method was used to obtain $\log_2$-transformed normalized expression values with the SCAN.UPC R package (version 2.26.0[301]), with annotation data from the Bioconductor project (version 3.9[174]). The microarray data generated from ABCF1 siRNA experiments is deposited in GEO with accession number GSE150541.

Analysis of processed microarray data

From the processed expression values, principal component analyses were performed with the 'prcomp' function (version 0.1.0) from the R statistical language (version 3.6.1; R Core Team, 2019) using default parameters. Determination of statistically significant differential gene expression was performed using the empirical Bayes method via the 'eBayes' function from the limma R package (version 3.40.0).[325] Statistical p values were adjusted using the Benjamini-Hochberg method, with a significance cutoff of 0.05. Significantly enriched Gene Ontology (GO) Biological Process Terms (ranked by p value) were determined using Enrichr ([82,221]). Scatter plots, PCA plots, and GO term enrichment dot plots were generated using the ggplot2 R package (version 3.2.1). Heat maps were generated using the pheatmap R package (version 1.0.12), with $\log_2$ expression scaled by gene and complete hierarchical clustering using a Euclidean distance measure applied. A GO term enrichment clustergram was modified from Enrichr using Inkscape.

Statistical analyses

All experiments were performed with an $n \geq 3$ unless otherwise noted. Experiments with HBEC-6KT and Calu-3 cells were considered independent when separated by a passage. Statistics were determined by permutation ANOVA with a Bonferroni-corrected *post-hoc* test comparing selected groups with $p < 0.05$ determined to be statistically significant.

2.2.3 Results

ABCF1 gene and protein expression is localized to human airway epithelial cells *in situ* and *in vitro*

Expression and functional studies of ABCF1 have focused on human synoviocytes, mouse embryonic fibroblasts, human embryonic kidney cells, and peripheral blood mononuclear cells.[25,294,295,319,388] I have demonstrated gene expression of ABCF1 in human airway epithelial cells.[11] To date, no *in situ* gene and protein expression data has confirmed ABCF1 expression in human lung tissues.

To address this knowledge gap, I first mined publicly available gene expression data from primary human airway epithelial cells from healthy subjects. *ABCF1* gene expression was observed along the airway generations (trachea, large, and small) at levels relative to *ABCC7/CFTR* and *ABCC4*, two other ABC transporters with reported functions in airway epithelial cells[90,132,181,323] (Figure 2.8 A). *CD34* and *CDH1* (encoding E-Cadherin) were used as negative and positive control genes, respectively, for airway epithelial cells to provide contextual expression levels.

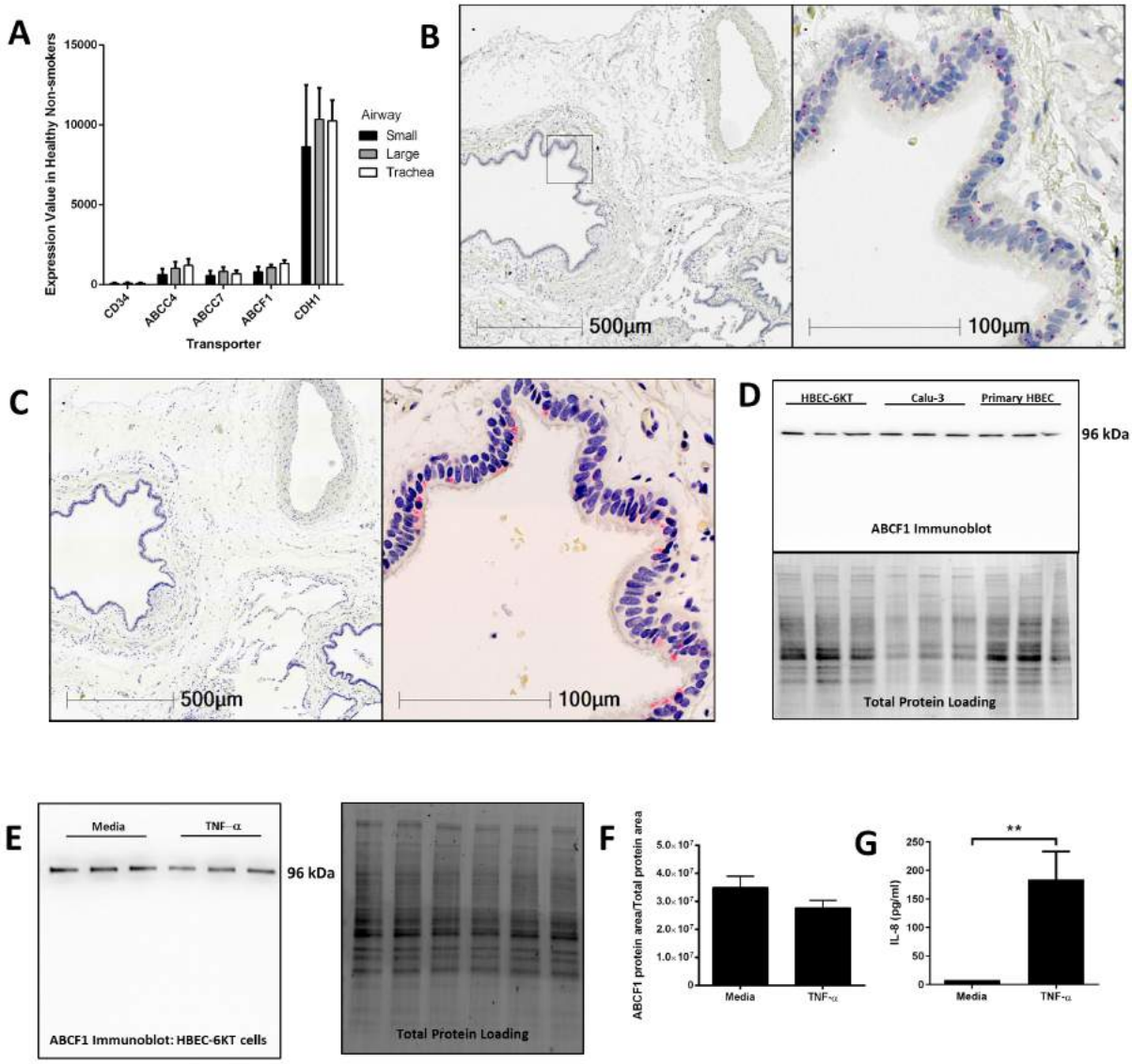

A
Expression Value in Healthy Non-smokers
15000
10000
5000
0
CD34
ABCC4
ABCC7
ABCF1
CDH1
Transporter
Airway
Small
Large
Trachea
B
500µm
100µm
C
500µm
100µm
D
HBEC-6KT
Calu-3
Primary HBEC
96 kDa
ABCF1 Immunoblot
Total Protein Loading
E
Media
TNF-α
96 kDa
ABCF1 Immunoblot: HBEC-6KT cells
Total Protein Loading
F
ABCF1 protein area/Total protein area
5.0x10⁷
4.0x10⁷
3.0x10⁷
2.0x10⁷
1.0x10⁷
0
Media
TNF-α
G
**
IL-8 (pg/ml)
250
200
150
100
50
0
Media
TNF-α

Figure 2.8: (A) Gene expression analysis of GEO deposited microarray data set (GSE11906) generated from epithelial cells isolated from trachea, large (2nd-3rd generation), and small airways (10th–12th generation) from healthy subjects (see Methods for details). (B) *In situ* hybridization of *ABCF1* RNAscope^{TM} probe in human lung under low (10X) and high (40X) magnification. Red puncta are representative of *ABCF1* gene transcripts with nuclei counter-stained blue. Representative image of $n = 10$. (C) Immunohistochemistry of ABCF1 in human lung under low (10X) and high (40X) magnification. Representative image of $n = 10$. Pink/red staining is representative of ABCF1 protein with nuclei counter-stained blue. (D) Immunoblot confirmation of ABCF1 protein expression in HBEC-6KT, Calu-3, and primary human airway epithelial cells (each cell type $n = 3$ distinct cell line passages or donors) with a single band observed at predicted molecular weight (96 kDa) with total protein loading blot demonstrating equal protein loading for each cell type. (E) Immunoblot of ABCF1 following TNF-α stimulation of HBEC-6KT cells with corresponding total protein stain. (F) Quantification of immunoblot of ABCF1 protein expression. (G) IL-8 secretion from HBEC-6KT cells measured by ELISA as positive control for TNF-α stimulation. All studies $n = 3$ unless otherwise noted. **: $p < 0.01$.

Next, *in situ* localization of *ABCF1* gene transcript was performed using RNAscope^{TM} probes on archived formalin fixed paraffin embedded human lung samples (Figure 2.8 B). *ABCF1* gene transcript was observed in small puncta throughout the cytoplasm and nuclear areas of airway epithelial cells. ABCF1 staining was also observed in submucosal cells with morphology consistent with macrophages.

Protein expression levels were next explored with validation of a commercially available antibody for ABCF1. Positive staining was observed in human airway epithelial cells as shown in a serial section used for in situ hybridization (Figure 2.8 C) with sparse staining in immune cells with macrophage morphology. Using *in vitro* culture of primary human airway epithelial cells and two distinct airway epithelial cell lines (Figure 2.8 D). For each airway epithelial cell type, a single band was observed at the predicted molecular weight of 96 kDa for ABCF1, validating the use of the antibody for *in situ* immunohistochemistry localization.

Lastly, to explore proposed regulatory mechanisms for ABCF1,[319] a TNF-α exposure was performed in human airway epithelial cells. Exposure to 10 ng/mL TNF-α for 24 hours failed to induce a change in ABCF1 protein expression (Figures 2.8 E, F), despite inducing an increase in IL-8 (Figure 2.8 G).

Collectively the *in vitro* and *in situ* data confirm gene and protein expression of ABCF1

in human airway epithelial cells, warranting downstream characterization and functional studies.

Reduced expression of ABCF1 under basal conditions *in vitro* has limited impact on HBEC-6KT viability and transcriptional profiles

Functional studies have implicated ABCF1 in translation initiation and have demonstrated that homozygous loss of function results in embryonic lethality.[294,295,388,413] Therefore, the basal functions of ABCF1 in human airway epithelial cells were interrogated in the context of cell viability and transcriptional profiling.

siRNA experiments were performed to reduce ABCF1 expression levels as no small molecule ABCF1 inhibitor has been described to date. Using siRNA approaches in HBEC-6KT, it is confirmed that ABCF1 protein levels can be reduced with qualitative (Figure 2.9 A) and quantitative measures (Figure 2.9 B). LDH levels were not elevated when ABCF1 expression was reduced with siRNA (Figure 2.9 C). Cell morphology was not different in human airway epithelial cells with reduced ABCF1 expression (Figure 2.9 D). Collectively, the quantitative and qualitative data suggest moderate levels of siRNA knockdown are not associated with compromised HBEC-6KT viability under basal conditions.

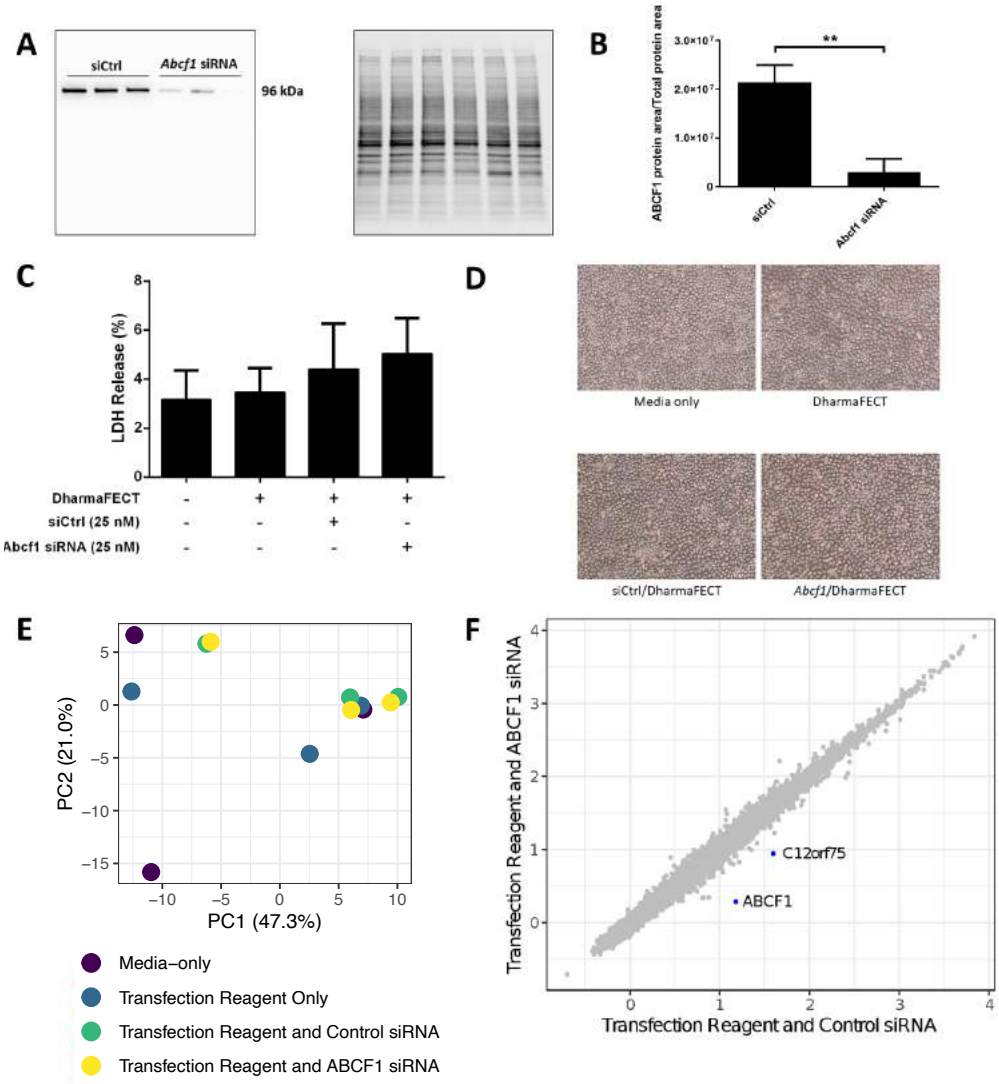

A
siCtrl
Abcf1 siRNA
96 kDa
B
ABCF1 protein area/Total protein area
3.0×10⁷
2.0×10⁷
1.0×10⁷
0
**
siCtrl
Abcf1 siRNA
C
LDH Release (%)
8
6
4
2
0
DharmaFECT - + + +
siCtrl (25 nM) - - + -
Abcf1 siRNA (25 nM) - - - +
D
Media only
DharmaFECT
siCtrl/DharmaFECT
Abcf1/DharmaFECT
E
PC2 (21.0%)
PC1 (47.3%)
Media−only
Transfection Reagent Only
Transfection Reagent and Control siRNA
Transfection Reagent and ABCF1 siRNA
F
Transfection Reagent and ABCF1 siRNA
Transfection Reagent and Control siRNA
C12orf75
ABCF1

Figure 2.9: (A) Immunoblot confirming siRNA-mediated knockdown of ABCF1 protein expression in HBEC-6KT cells. (B) Quantification of ABCF1 protein expression following siRNA treatment. (C) LDH quantification as a measure of cell viability following siRNA treatment. (D) Phase-contrast microscopy (4X magnification) of HBEC-6KT following siRNA treatment. (E) PCA plot of microarray gene expression profiles of HBEC-6KT cells following siRNA treatment. Purple circles (media alone), blue circles (transfection reagent only), green circles (transfection reagent and control siRNA), yellow circles (transfection reagent and ABCF1 siRNA). (F) Log_2 expression data for transfection reagent with ABCF1 siRNA compared to transfection reagent with control siRNA. Significantly differently expressed genes are in blue and are down-regulated (*ABCF1* and *C12orf75*). All studies $n = 3$; **: $p < 0.01$.

To interrogate the impact of reduced expression of ABCF1 under basal conditions, a human gene expression microarray analysis was performed. A principal component analysis of *ABCF1* reduction and corresponding experimental controls revealed no clustering between experimental replicates for any condition (Figure 2.9 E), suggesting that the overall impact of ABCF1 reduction under basal conditions minimally impacted global gene expression patterns. Statistical analysis comparing ABCF1 reduction and siRNA control treated HBEC-6KT confirmed *ABCF1* gene was down-regulated (Figure 2.9 F) which was associated with only one other significantly differentially expressed (up or down) gene, *C12orf75*, which encodes over-expressed in colon carcinoma-1 (OCC-1) protein.

Collectively these *in vitro* studies under basal conditions demonstrate that reduced *ABCF1* expression is not associated with changes in viability or significant genome wide changes in transcriptional profiles in HBEC-6KT.

The dsDNA viral mimic VACV-70 induces CXCL10 and an antiviral response in HBEC-6KT *in vitro*

Since reduced expression of ABCF1 under basal conditions resulted in limited impacts on cell viability and gene transcription, conditions of challenge in HBEC-6KT were next explored. ABCF1 was described as a dsDNA sensor in mouse embryonic fibroblasts that mediated CXCL10 secretion under challenge conditions with the viral mimic interferon stimulatory DNA (ISD) sequence,[228] a 45 bp oligomer shown to activate the STING-TBK1-IRF3 antiviral sensing axis.[186,367]

To determine the response of HBEC-6KT to ISD, a concentration-response study was performed followed by quantification of extracellular CXCL10 secretion (Figure 2.9 A).

ISD induced an increase in CXCL10 at 1 µg/mL while no increases were observed at lower (0.316 µg/mL) and higher (3.16 µg/mL) concentrations. Importantly, as concentration of ISD increased, the cellular response to the control (ssDNA of the ISD sequence) also increased. These results limited the use of ISD as dsDNA challenge stimulus in HBEC-6KT cells for studying ABCF1 function.

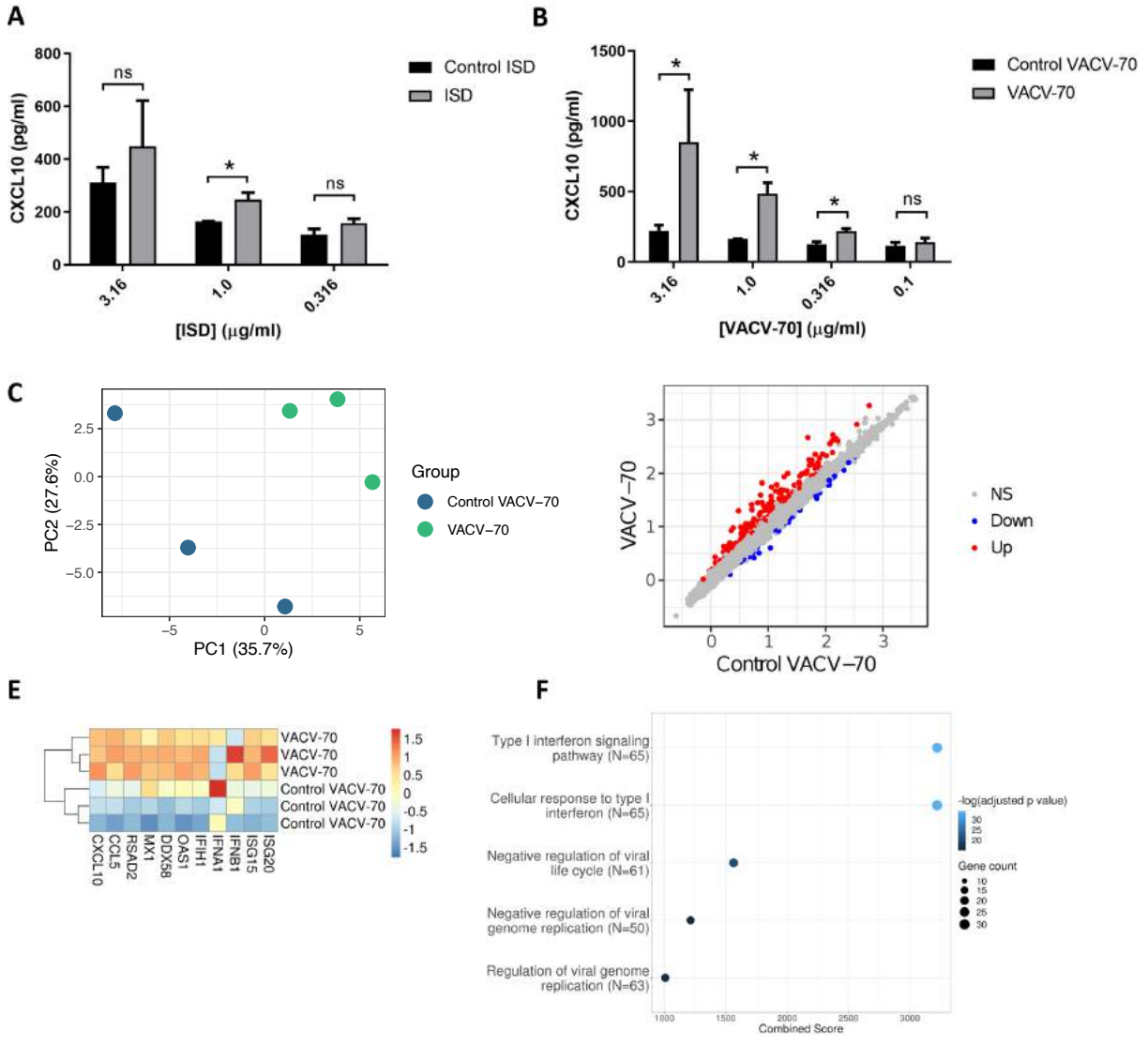


Figure 2.10: (A) Concentration-response analysis of ISD-induced CXCL10 protein production for HBEC-6KT cells. Gray bars: ISD, Black bars: control ssDNA generated from ISD. (B) Concentration-response analysis of VACV-70-induced CXCL10 protein production for HBEC-6KT cells. Gray bars: VACV-70, Black bars: control ssDNA generated from VACV-70. (C) PCA plot of microarray gene expression profiles of HBEC-6KT cells following transfection with VACV-70 or control. Blue circles (control VACV-70), green circles (VACV-70). (D) Log_2 expression data for transfection treatment with VACV-70 compared to control VACV-70. Significantly differently expressed genes are identified in blue (down: 42 genes) and red (up: 170 genes). (E) Heat map of log_2 expression data (scaled by gene) of select known antiviral for VACV-70 and control VACV-70 samples. (F) Top 5 GO Biological Processes are ranked by increasing $-log_{10}$ adjusted p value, with number (Count) of significantly differentially expressed genes between VACV-70 and control VACV-70 contributing to the total number of genes associated with the given pathway (N) denoted by the size of circle. All studies $n = 3$; *: $p < 0.05$.

Vaccinia virus is a dsDNA virus that is able to infect airway epithelial cells.[18,56,100,318,397] The response of HBEC-6KT to VACV-70, a 70 bp dsDNA oligonucleotide containing Vaccinia virus motifs[390] was therefore determined. VACV-70 induced a concentration dependent increase in CXCL10 from 0.316 µg/mL to 3.16 µg/mL. In contrast to ISD, no cellular response to the control (ssDNA of the VACV-70 sequence) was observed at any concentration, based on ANOVA with Bonferroni correction.

To characterize the molecular pathways activated by VACV-70, a transcriptional and pathway analysis of HBEC-6KT was performed following challenge. To interrogate the VACV-70 transcriptional responses a principal component analysis was performed for microarray gene expression data, revealing distinct clustering between stimulation (VACV-70) and control (Figure 2.10 C). Statistical analysis revealed 170 up-regulated genes and 42 down-regulated genes with VACV-70 stimulus (Figure 2.10 D). VACV-70 up-regulated *CXCL10* gene expression and a curated list of antiviral related interferon stimulated genes (Figure 2.10 E). GO term analysis revealed that the top pathways activated by VACV-70 were associated with type I interferon signaling, viral responses, and cellular responses to viruses (Figure 2.10 F).

Collectively these *in vitro* challenge studies confirm that VACV-70, a dsDNA viral mimic, can induce CXCL10 and antiviral transcriptional responses in HBEC-6KT.

Reduced expression of ABCF1 under VACV-70 stimulated conditions attenuates CXCL10 secretion with limited impact on HBEC-6KT viability

VACV-70 induction of CXCL10 in HBEC-6KT at the gene (Figure 2.10 E) and protein (Figure 2.10 B) levels have been confirmed. Furthermore, it has been demonstrated that siRNA-mediated reduction of ABCF1 expression is possible and has no impact on cell viability (Figures 2.9 A–D). Therefore, a VACV-70 challenge with reduced expression of ABCF1 by siRNA was next performed with a primary readout of CXCL10.

ABCF1 reduction was associated with a decrease in CXCL10 protein secretion from HBEC-6KT, with confirmation and quantification of ABCF1 reduction performed by immunoblot (Figures 2.11 A–C). Cell viability following VACV-70 challenge and ABCF1 attenuation was not impacted as assessed by LDH quantification (Figure 2.11 D). Qualitative analysis following VACV-70 challenge and ABCF1 reduction revealed no impact on HBEC-6KT cell morphology (Figure 2.11 E).

Collectively these *in vitro* challenge and functional studies demonstrate that ABCF1 siRNA treatment attenuated VACV-70-induced CXCL10 protein secretion in HBEC-6KT.

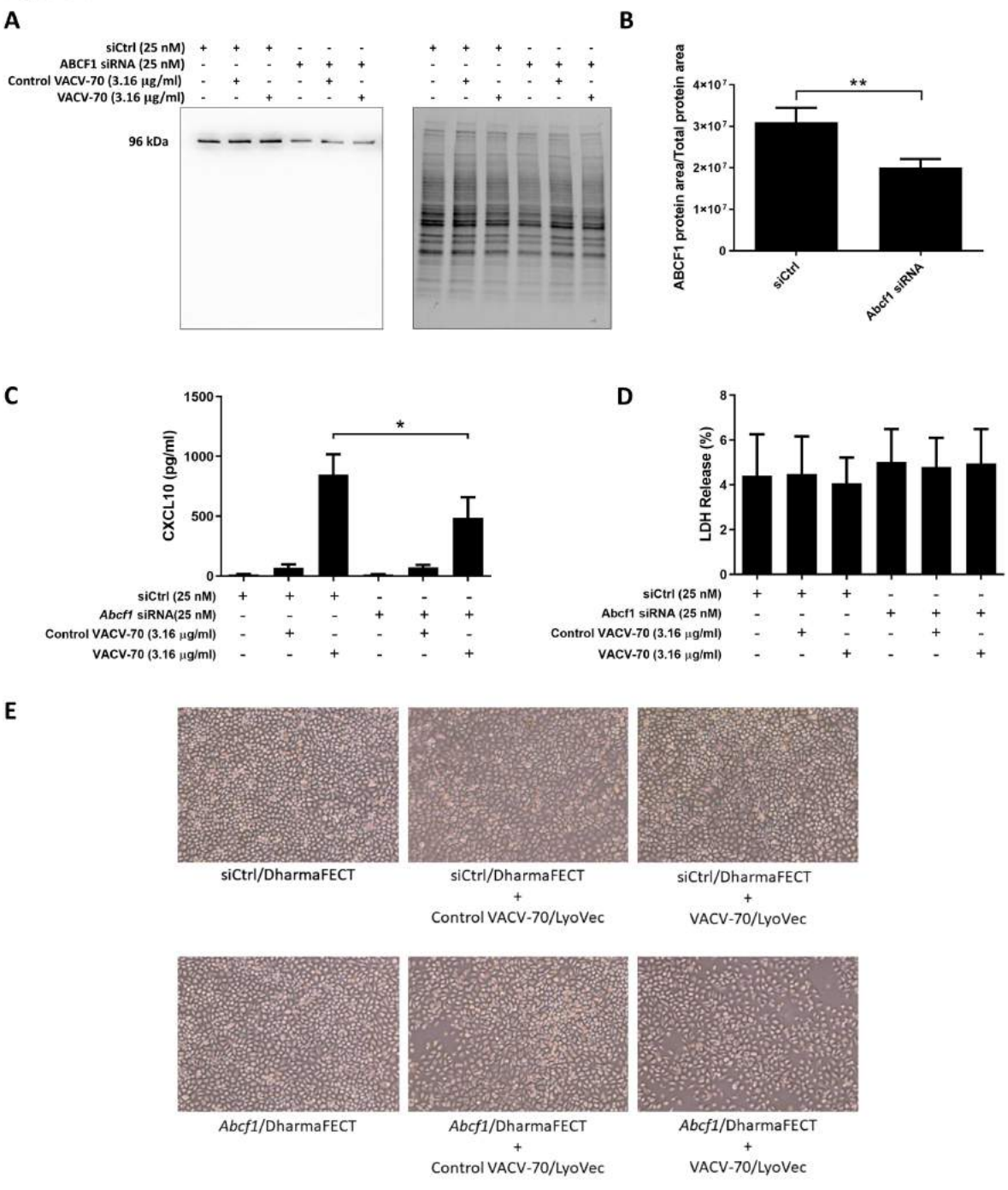

A
siCtrl (25 nM)
ABCF1 siRNA (25 nM)
Control VACV-70 (3.16 µg/ml)
VACV-70 (3.16 µg/ml)
96 kDa
B
ABCF1 protein area/Total protein area
4×10^7
3×10^7
2×10^7
1×10^7
0
**
siCtrl
Abcf1 siRNA
C
CXCL10 (pg/ml)
1500
1000
500
0
*
siCtrl (25 nM)
Abcf1 siRNA(25 nM)
Control VACV-70 (3.16 µg/ml)
VACV-70 (3.16 µg/ml)
D
LDH Release (%)
8
6
4
2
0
siCtrl (25 nM)
Abcf1 siRNA (25 nM)
Control VACV-70 (3.16 µg/ml)
VACV-70 (3.16 µg/ml)
E
siCtrl/DharmaFECT
siCtrl/DharmaFECT + Control VACV-70/LyoVec
siCtrl/DharmaFECT + VACV-70/LyoVec
Abcf1/DharmaFECT
Abcf1/DharmaFECT + Control VACV-70/LyoVec
Abcf1/DharmaFECT + VACV-70/LyoVec

Figure 2.11: (A) Immunoblot confirming siRNA-mediated knockdown of ABCF1 protein expression in HBEC-6KT cells under experimental conditions of VACV-70 challenge. (B) Quantification of ABCF1 protein expression following siRNA treatment. (C) VACV-70 (3.16 µg/mL)-induced CXCL10 protein production for HBEC-6KT cells with ABCF1 siRNA treatment. Black bars: siCtrl treated. Gray bars: ABCF1 siRNA treated. (D) LDH quantification as a measure of cell viability following VACV-70 and siRNA treatment. (E) Phase-contrast microscopy (4X magnification) of HBEC-6KT following VACV-70 and siRNA treatment. All studies $n = 3$; *: $p < 0.05$ and **: $p < 0.01$.

ABCF1 reduction does not impact VACV-70-induced antiviral transcriptional responses in HBEC-6KT

In parallel to induction of *CXCL10* gene, GO pathway analysis confirms that VACV-70 induces a dominant antiviral transcriptional signature (Figure 2.10 E, F). Therefore, how reduced expression of ABCF1 impacts transcriptional responses induced by VACV-70, beyond induction of CXCL10, was next explored. A principal component analysis was performed for microarray gene expression data, revealing distinct clustering in samples where ABCF1 expression was reduced relative to control conditions under conditions of VACV-70 challenge (Figure 2.13 A; yellow versus blue). Statistical analysis revealed 63 up-regulated genes and 65 down-regulated genes when comparing ABCF1 reduction relative to control under conditions of VACV-70 challenge (Figure 2.13 B). siRNA mediated reduction of ABCF1 was confirmed and associated with attenuation of *CXCL10* gene expression (Figure 2.13 C, $p = 0.06$).

To explore a focused transcriptional response of ABCF1 reduction in the context of VACV-70 challenge, a hypothesis-directed approach curated 79 genes from the GO term "Regulation of defense response to virus" and key components of viral sensing for heat map visualization[138] (Figure 2.13 D). Statistical analysis revealed no global significant difference between ABCF1 reduction and control groups for the expression pattern of this curated list of genes. To explore the broader transcriptional responses of ABCF1 reduction in the context of VACV-70 challenge, a data-directed approach with GO term analysis was performed. Top-ranking GO pathway terms included "Regulation of toll-like receptor 3–4 signaling pathways", which were driven by the genes *WDFY1*, *TNFAIP3*, and *NR1D1* (Figure 2.13 E, F). Complementary cytokine analysis further revealed impacts of ABCF1 reduction on PDGF-BB, VEGF-A, and to a lesser extent IL-6, IL-8, and IL-1 family members, IL-1a, IL-18, and IL-1RA (Figure 2.12).

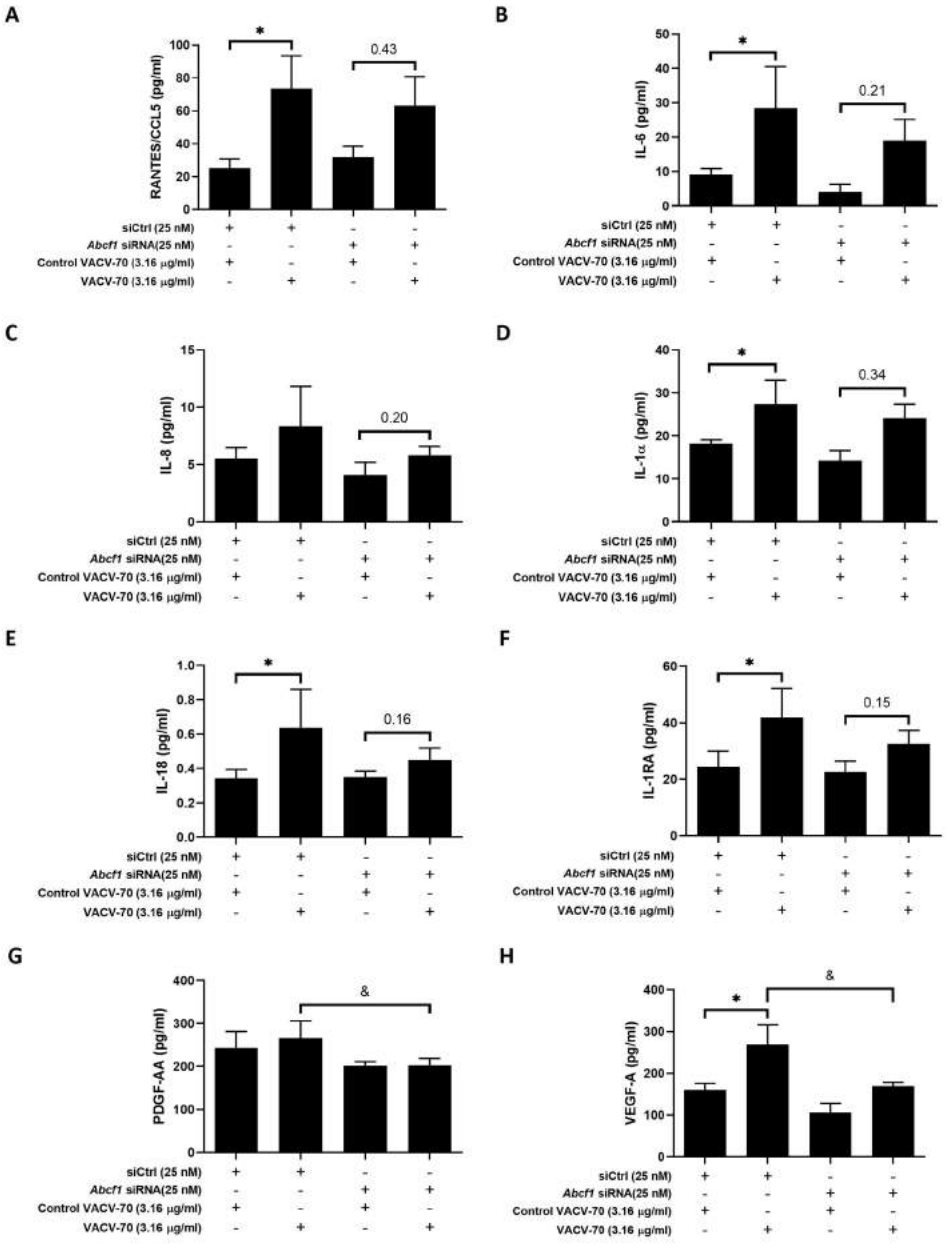


Figure 2.12: VACV-70 (3.16 µg/mL)-induced cytokine protein production for HBEC-6KT cells with ABCF1 siRNA treatment. (A) RANTES/CCL5, (B) IL-6, (C) IL-8, (D) IL-1a, (E) IL-18, (F) IL-1RA, (G) PDGF-AA, (H) VEGF-A. All studies $n = 3$; *: $p < 0.05$ relative to siCtrl + ctrl VACV-70; &: $p < 0.05$ relative to siCtrl + VACV-70.

As the data suggested that ABCF1 functions in HBEC-6KT may extend beyond sensing of VACV-70 dsDNA viral mimic through regulation of TLR signaling, Poly I:C, a dsRNA analog and TLR3 ligand that induces interferon responses including CXCL10 production, was explored. ABCF1 reduction was associated with a 63% reduction in Poly I:C-induced CXCL10 protein ($p = 0.07$) secretion but did not impact Poly I:C-induced IL-8 protein secretion (Figure 2.13 G, H).

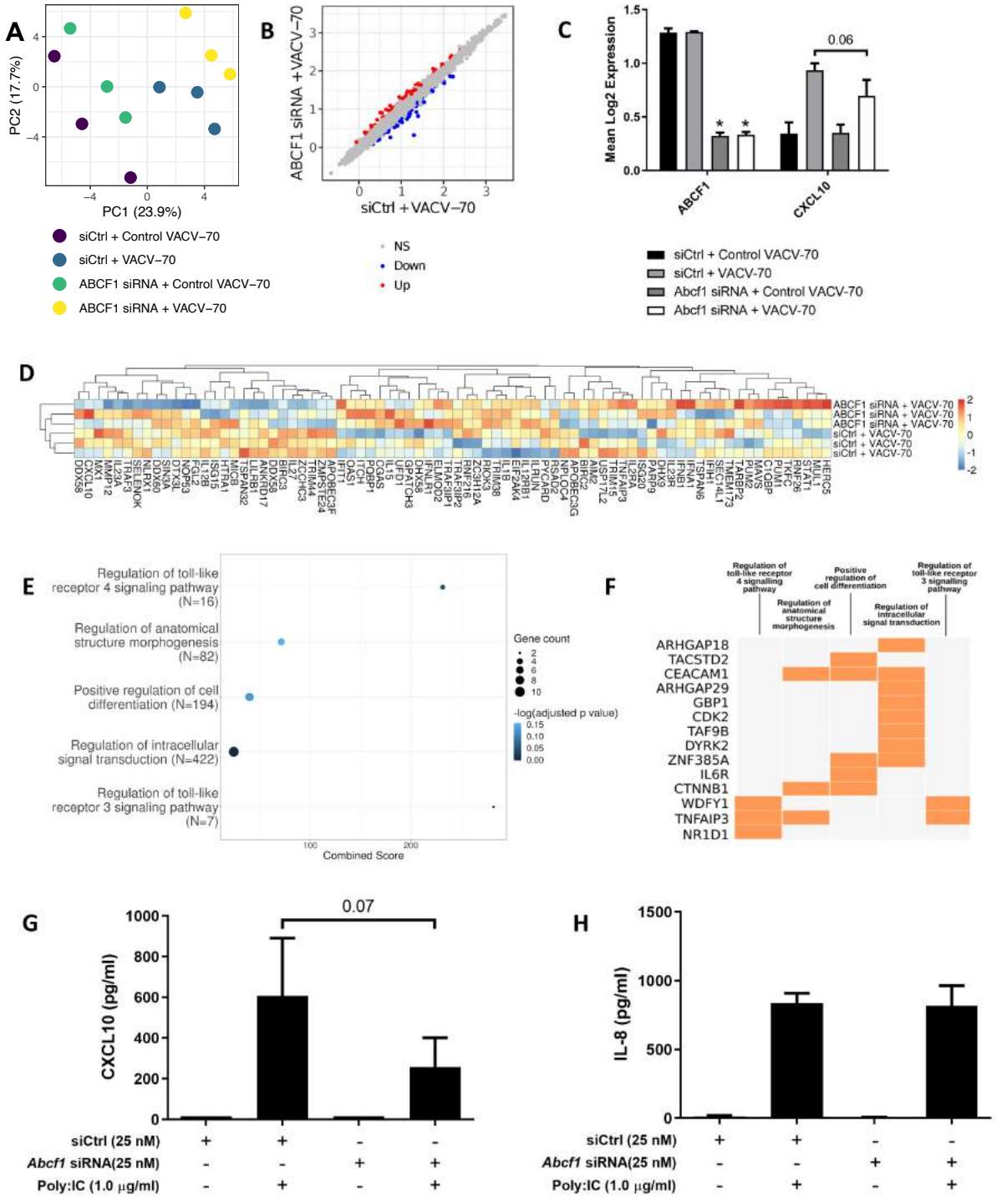

A
PC2 (17.7%)
PC1 (23.9%)
siCtrl + Control VACV−70
siCtrl + VACV−70
ABCF1 siRNA + Control VACV−70
ABCF1 siRNA + VACV−70
B
ABCF1 siRNA +VACV−70
siCtrl +VACV−70
NS
Down
Up
C
Mean Log2 Expression
0.06
ABCF1
CXCL10
siCtrl + Control VACV-70
siCtrl + VACV-70
Abcf1 siRNA + Control VACV-70
Abcf1 siRNA + VACV-70
D
ABCF1 siRNA + VACV-70
ABCF1 siRNA + VACV-70
ABCF1 siRNA + VACV-70
siCtrl + VACV-70
siCtrl + VACV-70
siCtrl + VACV-70
E
Regulation of toll-like receptor 4 signaling pathway (N=16)
Regulation of anatomical structure morphogenesis (N=82)
Positive regulation of cell differentiation (N=194)
Regulation of intracellular signal transduction (N=422)
Regulation of toll-like receptor 3 signaling pathway (N=7)
Combined Score
Gene count
-log(adjusted p value)
F
Regulation of toll-like receptor 4 signalling pathway
Positive regulation of cell differentiation
Regulation of toll-like receptor 3 signalling pathway
Regulation of anatomical structure morphogenesis
Regulation of intracellular signal transduction
ARHGAP18
TACSTD2
CEACAM1
ARHGAP29
GBP1
CDK2
TAF9B
DYRK2
ZNF385A
IL6R
CTNNB1
WDFY1
TNFAIP3
NR1D1
G
CXCL10 (pg/ml)
0.07
siCtrl (25 nM)
Abcf1 siRNA(25 nM)
Poly:IC (1.0 μg/ml)
H
IL-8 (pg/ml)
siCtrl (25 nM)
Abcf1 siRNA(25 nM)
Poly:IC (1.0 μg/ml)

Figure 2.13: (A) PCA plot of microarray gene expression profiles of HBEC-6KT cells following ABCF1 knockdown and VACV-70 treatment. Purple circles (control siRNA and control VACV-70), blue circles (control siRNA and VACV-70), green circles (ABCF1 siRNA and control VACV-70), yellow circles (ABCF1 siRNA and VACV-70). (B) Log_2 expression data for transfection treatment with ABCF1 siRNA and VACV-70 compared to control siRNA and VACV-70. Significantly differently expressed genes are identified in blue (down: 65 genes) and red (up: 63 genes). (C) Confirmation of ABCF1 and CXCL10 attenuation with ABCF1 siRNA treatment presented as $\log_2$ expression data. Black bars: control siRNA and control VACV-70, light gray bars: control siRNA and VACV-70, dark gray bars: ABCF1 siRNA and control VACV-70, white bars: ABCF1 siRNA and VACV-70. (D) Heat map of $\log_2$ expression data (scaled by gene) of genes associated with the "regulation of defense response to virus" GO term (68 genes) plus the selected known antiviral genes from Figure 2.10 E (11 genes) for VACV-70 samples with and without ABCF1 siRNA. (E) Top 5 GO Biological Processes are ranked by increasing -$\log_{10}$ adjusted p value, with number (Count) of significantly differentially expressed genes between VACV-70 samples with and without ABCF1 siRNA, contributing to the total number of genes associated with the given pathway (N) denoted by the size of circle. (F) Significantly differentially expressed genes between VACV-70 samples with and without ABCF1 were submitted to Enrichr for generation of a clustergram defining the gene contribution (y axis) to the functional enrichment of the top 5 GO Biological Processes (x axis), with orange squares denoting the association of a differentially expressed gene with a particular GO term. (G) Poly I:C (1.0 µg/mL)-induced CXCL10 and (H) IL-8 secretion for HBEC-6KT cells with ABCF1 siRNA treatment. All studies $n = 3$; *: $p < 0.05$.

Collectively these *in vitro* challenge and functional studies with transcriptional analyses demonstrate a role for ABCF1 in mediating VACV-70 and Poly I:C induced CXCL10 secretion and TLR3 related signaling in HBEC-6KT.

2.2.4 Discussion

The human airway epithelium expresses a variety of sensors that can detect and initiate an immune response to virus infection. Recognition by these sensors can trigger downstream activation of antiviral responses by inducing the production and release of antiviral and inflammatory cytokines.[204] This includes TLRs found at the cell surface such as TLR3, which can detect viral RNA to trigger a type I IFN response by the TRIF signaling pathway. In addition, TLR7 and TLR8 detect viral RNA while TLR9 detects CpG containing

viral DNA in the endosomes triggering the same response through the MyD88 signaling pathway.[204] In addition to TLRs, there are several cytosolic receptors including RIG-I-like receptors that recognizes viral RNA, as well as cytosolic DNA sensors such as cyclic GMP-AMP synthase, AIM2-like proteins and DNA-dependent activator of IFN-regulatory factors that produces an IFN response.[138] It is likely that additional candidates are present as redundancy is built into viral sensing mechanisms in host cells.[274]

ABCF1, a member of the ABC transporter family expressed in diverse mammals and different tissue types, has been reported to have diverse functions including initiation of mRNA translation, dsDNA viral sensing, and polarization of immune cell phenotype.[25,228,294,295,319,388] I have recently reported *ABCF1* gene expression levels in human airway epithelium,[11] but the function of this molecule remained unexplored. Herein I confirm ABCF1 gene and protein expression *in situ* and *in vitro* in primary human lung tissue and cell lines and explore its function in airway epithelial cells. Under basal conditions, reduced expression of ABCF1 did not lead to quantitative changes in cell viability or qualitative changes in cell morphology associated with cell death. Furthermore, ABCF1 reduction failed to significantly alter basal transcriptional activity in a human airway epithelial cell line, HBEC-6KT. Under VACV-70 challenge, a model of dsDNA viral exposure, ABCF1 was linked to CXCL10 secretion. Interestingly, despite the demonstrated activation of a viral gene signature by VACV-70, no global change in antiviral gene expression patterns were observed with ABCF1 reduction. In contrast, the gene pathways regulated by ABCF1 under VACV-70 challenge were associated with TLR signaling and intracellular signal transduction. Furthermore, Poly I:C, a dsRNA analog and TLR3 ligand induced CXCL10 in an ABCF1 sensitive mechanism. Collectively, this study's data suggests that ABCF1 may regulate CXCL10 production downstream of dsDNA sensing mechanisms and TLR3 in human airway epithelial cells. It remains possible that ABCF1 can function to complement viral sensing mechanisms mediated by canonical dsRNA viral response machinery (e.g., RIG-I) as there are possible redundancies in viral sensing mechanisms in the cell.[274]

ABCF1 (originally called ABC50) was first identified in human synoviocytes at the mRNA level as a transcript regulated by TNF-α exposure.[319] ABCF1 is unique in the mammalian ABC transporter family in that it contains the signature ATP binding LSGGQ amino acid motif and associated Walker A and B motifs for phosphate binding, but lacks a predicted transmembrane region,[160,242,319] which is supportive of a cytosolic localization and function. *ABCF1* transcript expression profiling has revealed near ubiquitous expression in human organs including lung, heart, brain, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocytes.[319] The expression of ABCF1 has subsequently been identified in the human HeLa cells and embryonic kidney cells and other mammalian cells from rats, rabbits, ham-

sters, and mice.[25,294,295,388,413] Highlighting the importance of ABCF1 in normal physiology and development, homozygous deletion of ABCF1 is embryonic lethal in either C57Bl/6 mice or BALB/c mice.[413] As I recently identified gene expression of *ABCF1* in human airway epithelial cells,[11] I set out to first confirm this at the protein level and then determine the function(s) of ABCF1 in human airway epithelial cells. It is confirmed that *ABCF1* gene expression is present in airway epithelial cells and expressed at levels relative to other known ABC transporters with function in this cell type, *ABCC4* and *ABCC7/CFTR*.[90,132,180,323] *In situ* hybridization using RNAscope™ technology demonstrated *ABCF1* transcripts present in the airway epithelial cells in human lung samples, which was consistent with positive immunohistochemical staining of protein in a serial section of the same samples using an antibody validated for specificity. Since the original discovery of ABCF1 was the result of an up-regulated transcript resulting from TNF-α stimulation of synoviocytes, it was examined if this mechanism was functional in human airway epithelial cells. In contrast to the reported data on synoviocytes, TNF-α stimulation failed to induce ABCF1 protein expression in HBEC-6KT, despite IL-8 induction as a positive control. Interestingly, in a recent report profiling the role of ABCF1 in murine bone-marrow derived macrophages, TNF-α stimulation suppressed ABCF1 protein expression.[25] These contrasting observations of ABCF1 expression in HBEC-6KT compared to synoviocytes and bone-marrow derived macrophages could be due to the difference in ontogeny and function of these cells. Collectively these results and those in the literature support gene and protein expression of ABCF1 in human airway epithelial cells, and that regulation of this protein is likely to be cell specific.

The first description of a potential function for ABCF1 in mammalian cells was derived from the experiments on human synoviocytes, suggesting a role in translation due to homology of molecular sequence with yeast proteins that performed this function.[319,388] The embryonic lethality observed in mice for homozygous ABCF1 deletion and ubiquitous expression across multiple cell and tissue types,[413] is consistent with ABCF1 playing a role in a fundamental biological process like protein translation. The observation that proliferating cells including synoviocytes stimulated with TNF-α and T cells stimulated with phorbol myristate acetate and ionomycin, elevate ABCF1 levels is further consistent with a role in translation.[318,388] Subsequent to the discovery of ABCF1 gene expression and homology modeling, biochemical studies implicated the protein in interaction with eukaryotic initiation factor-2 (eIF2), a heterotrimeric protein consisting of α, β, and γ subunits, that is important for translation initiation.[388] A distinguishing feature of ABCF1 relative to other ABC transporters is a N-terminal domain that is able to interact with eIF2α in a process that potentiates binding of methionyl-tRNA and initiation of translation.[294] In addition to eIF2α interactions, ABCF1 associates with ribosomes in a process potentiated

by ATP binding to the nucleotide binding domains and inhibited by ADP,[388] although the hydrolysis of ATP seems dispensable for ribosome interaction.[294] To explore the potential function of ABCF1 as an initiator of translation in human airway epithelial cells, a siRNA approach was undertaken to attenuate gene and protein expression followed by a global assessment of cell viability and transcriptomics. Surprisingly, under basal conditions, reduced expression of ABCF1 at the gene and protein level did not impact HBEC-6KT viability, morphology or transcriptional profile. Importantly, the outcome measurements were performed on HBEC-6KT that were sub-confluent and undergoing proliferation in serum-free media, an experimental condition where ABCF1 function in translation initiation would be relevant. A limitation of the study design is that global gene expression was measured under the assumption that this would reflect any global changes in gene translation, an indirect approach which does not allow us to directly implicate ABCF1 expression levels to protein synthesis. Interestingly, the observations of minimal changes in HBEC-6KT may be consistent with cells of epithelial lineage, as near complete ABCF1 knockdown in HeLa cells was also only associated with a modest attenuation of total protein synthesis.[295] Collectively, these results suggest that ABCF1 may function independent of protein translation functions in HBEC-6KT, as gene and protein attenuation results in no changes in cell viability or global transcriptional profile.

The original discovery that ABCF1 expression was regulated by TNF-α stimulation suggested a link to immune responses, although no differential expression patterns were observed for synoviocytes from healthy individuals or those with rheumatoid arthritis.[319] Subsequently, ABCF1 has been implicated in immune responses via a cytosolic dsDNA viral sensing function using mouse embryonic fibroblasts.[228] Using an integrative bioinformatic and molecular biology approach, a biotinylated ISD sequence was used as a bait and transfected into cells, followed by proteomic interrogation of identified candidates. The ISD bait method was validated by identifying known dsDNA sensors including HMGB1, HMGB2, and HMGB3, components of the AIM2 inflammasome, and the SET complex that plays a role in HIV-1 retroviral detection and infection.[426] Within the pool of unknown dsDNA interacting candidates, ABCF1 was mechanistically linked to ISD induced-CXCL10 production using siRNA methods. The observed ISD induced-CXCL10 converged on IRF3 signaling, confirmed by showing reduced IRF3 phosphorylation following ISD treatment under conditions of ABCF1 silencing. In a separate study, ABCF1 has been implicated as a molecular switch downstream of TLR4 signaling in mouse bone-marrow derived macrophages that regulates MyD88 dependent pro-inflammatory and TRIF/TRAM dependent anti-inflammatory processes.[25] Using *in vitro* an *in vivo* model systems, ABCF1 was implicated in polarizing pro-inflammatory macrophages to an anti-inflammatory/tolerant macrophage phenotype with direct involvement in shifting the systemic inflammatory response syndrome

to a endotoxin tolerance phase in sepsis.[25] The mechanism responsible for the ABCF1-mediated polarization of macrophages was identified to be a E2-ubiqutin-conjugating enzyme function. In wild-type macrophages the TRIF-IFN-β pathway is intact with attenuation of the MyD88 pathway, allowing IRF3 phosphorylation, dimerization, and IFN-β expression. In contrast, heterozygosity for ABCF1 results in attenuation of the TRIF-IFN-β pathway, with reduced IRF3 activation and IFN-*beta* production. Importantly, these two immunological studies converge on a relationship between ABCF1 and IRF3, which could involve direct or indirect interactions to facilitate downstream immune responses. Consistent with the potential role for ABCF1 as a dsDNA sensor, immune and transcriptional responses were explored downstream of VACV-70, a dsDNA viral mimic capable of activating STING, TBK1, and IRF3 independent of TLRs.[390] VACV-70 induced a dominant antiviral signature and pathway activation in HBEC-6KT, consistent with successful transfection and cytosolic sensing. Reduced expression of ABCF1 was associated with a reduction in CXCL10, an antiviral cytokine regulated by IRF3 activation, independent of any changes in cell viability or morphology. Transcriptomics revealed that although attenuation of CXCL10 was observed with ABCF1 siRNA, a global attenuation of an antiviral signature was not observed. GO analysis identified that the key pathways that were significantly impacted by ABCF1 siRNA treatment during VACV-70 challenge were related to TLR signaling. Interestingly, a key gene identified in the VACV-70 challenge and ABCF1 silencing studies is *WDFY1*, which links TLR3/4, TRIF, and IRF3 signaling.[173] This finding suggested that ABCF1 could potentially be regulating both TLR4 and TLR3/TRIF/IRF3 signaling.[25] This hypothesis was tested by using Poly I:C, a dsRNA viral mimic that activates TLR3 and IRF3.[138] ABCF1 siRNA treatment attenuated Poly I:C-induced CXCL10 production, further demonstrating a link between ABCF1 and TRIF/IRF3, perhaps through regulation of *WDYF1*. While the link between ABCF1 and TLR4 has not yet been demonstrated, a LPS challenge with ABCF1 siRNA treatment would effectively interrogate this. These exploratory results suggest that ABCF1 is likely to play a complex role in innate immunity in response to cytosolic nucleic acids, with a potential interaction with TRIF/IRF3 for regulation of CXCL10.

Throughout this study several technical issues were encountered. The absence of pharmacological interventions that could antagonize ABCF1 function required us to pursue molecular approaches with siRNA. siRNA approaches were unable to completely attenuate ABCF1 at concentrations of 25 nM for up to 48 hours. Longer durations of silencing were not possible as the human airway epithelial cell line used showed changes in morphology with vehicle control transfection reagent beyond 48 hours of incubation. The inability to completely attenuate ABCF1 levels was consistent with human embryonic kidney cells.[295] Secondary to addressing ABCF1 expression levels, this study sought to explore the functional

consequences with the reported dsDNA viral mimic ISD as reported in the literature with mouse embryonic fibroblasts.[25] During the concentration-response studies with ISD, the vehicle control condition resulted in elevations in the primary readout of CXCL10, which suggested an unexplained confounding factor. This study therefore opts to use VACV-70 in place of ISD, which limits the ability to directly compare these results to those that have established ABCF1 as a dsDNA sensor with ISD.[25] Importantly, reported findings from the HBEC-6KT cell line will require confirmation using primary human airway epithelial cells under submerged and/or air-liquid interface culture conditions to more accurately model in the *in situ* human respiratory mucosa.

In conclusion, this study confirms that ABCF1 is expressed at the gene and protein level *in situ* and *in vitro* in human airway epithelial cells. In HBEC-6KT, ABCF1 has minimal functions for cell viability and transcriptional regulation under basal conditions but is important for mediating immune responses to cytosolic nucleic acids in pathways that involve TLR signaling and CXCL10 production. The data form the foundation to pursue precisely how ABCF1 is regulated and where it functions in the network of cytosolic nucleic acid sensors and immune responses in human airway epithelial cells.

2.3 Summary

The goal of this chapter was to leverage bioinformatic techniques to interrogate the effects of well-characterized airway particulates, thereby elucidating previously unexplored outcomes of these inhaled stimuli on the human lung epithelium. The studies described in this chapter demonstrate the myriad functions of the ABC transporter family within the context of the human airway, from maintaining surfactant homeostasis to potential viral sensing, while simultaneously further characterizing the wide-reaching effects of ubiquitous airway particulates on the human lung.

Through targeted statistical analysis, I was able to identify important, yet previously unobserved functional outcomes related to ABC transporters. Well-annotated, publicly available data sets were the foundation of these studies and are excellent resources to explore *a priori* hypotheses as they provide large enough samples sizes to confer necessary statistical power. These case studies reinforce the concept that novelty can come even from previously well-characterized sources by utilizing informatics. This raises important questions: what are the capabilities of current computational methods when performing exploratory, data-directed analyses and how can these methods be harnessed to understand other relevant yet under-studied or completely novel airway exposures?

Chapter 3

Cannabis smoke exposure

Cannabis is the most commonly consumed illicit drug worldwide, with the United Nations World Drug Report estimating over 180 million global cannabis users.[279] Decriminalization and legalization of medicinal and/or recreational cannabis in several global jurisdictions have led to increased access to cannabis products that will likely lead to increased use over the coming years. Canada legalized cannabis in 2017 and concomitantly initiated the annual Canadian Cannabis Survey to monitor perception and use patterns. These surveys demonstrate that 20% of the general population consumes cannabis[68–70] with a subset (12%) of the population using cannabis for medicinal purposes. In Canada, combustion is the most common form of consumption as reported by 94 and 89% of users in the 2017 and 2018 Canadian Cannabis Surveys, respectively.[68,69] With over 90% of cannabis consumers having identified combusted smoke inhalation as a route of delivery, the lung can be highlighted as a dominant target for cannabis exposures. Within the cannabis consumer population, there is a skewed distribution of use patterns, with males identified as consuming more frequently in the past 12 months relative to females (26% versus 18%). The prospective design of the Canadian Cannabis Surveys has demonstrated that cannabis use over time is widely accepted at a population level for recreational and medicinal purposes with no signal for decline in use post-legalization.[68–70] Additionally, the data demonstrate a strong preference for inhalation routes of delivery.

Inhalation of cannabis smoke delivers pharmacologically active ingredients to the lung, including tetrahydrocannabinol (THC) and cannabidiol (CBD), as well as combustion products such as polycyclic aromatic hydrocarbons.[266] Many studies have investigated the psychoactive and immunomodulatory properties of THC and CBD respectively, and examined the potential use of cannabis as an intervention for chronic pain, immune disorders and neurological disorders.[6,201] Despite this body of knowledge and recent rise of systematic

reviews on these topics, the effects of inhaled cannabis smoke on the respiratory mucosal immune responses are less clear.[317]

The negative effects of tobacco smoke on the lung and its airway epithelium are universally accepted.[27,193,381] Tobacco smoke induces oxidative stress and cytokine production while disrupting epithelial barrier function, impacting chronic lung disease pathology and exacerbations.[362] In contrast, there remains a paucity of observational and mechanistic cannabis smoke exposure studies and models[373] that can inform our understanding of potential public health issues. Early clinical exposure studies demonstrated that acute cannabis smoke exposure in healthy subjects is able to provide a sustained increase in lung function, which contrasted tobacco smoke inhalation.[376,391] In asthmatic subjects, cannabis smoking is able to produce a rapid reversal of exercise-induced bronchoconstriction and minimize bronchoconstriction induced by methacholine inhalation.[377,378] Despite these objective benefits in a controlled laboratory setting, critical consideration needs to be given to the observed negative impacts on lung function in the context of increased frequency and intensity of cannabis smoking. Contrasting acute cannabis exposure, cannabis smoking over a two-month period was associated with a decrease in airway compliance that was correlated with quantity of cannabis consumed.[379] Like tobacco smoke, repeated exposure to cannabis smoke has been associated with cough and shortness of breath relative to non-smokers.[17,371,374,420] Further consolidating the negative impacts of chronic cannabis use on lung health, population level analyses reveal that greater intensity of cannabis smoking is correlated with reduced lung function and increased risk of developing chronic obstructive pulmonary disease (COPD).[370,371] Furthermore, in multiple independent cohorts, cannabis smoking has been associated with a pro-inflammatory phenotype in the lung, associated with bronchitis and impaired immune cell function.[32,117,134,150,326] However, few studies have examined the effects of cannabis smoke on the respiratory mucosa, the primary physical and immunological barrier to the environment.

The lungs are in constant contact with harmful environmental agents such as viruses and bacteria yet we rarely show signs of infection.[179,288] Minimized infection is the result of innate and adaptive immune processes that include the physical barrier and immunological functions of the airway epithelium and mucosa. Protection rendered by the airway epithelium can be compromised by tobacco smoke,[19,111,177,320] leading to increased susceptibility to infections and potential for host pathology. Whether cannabis smoke exposure similarly impacts airway epithelial cell function and immunity relevant in pathogen defence remains to be determined.

Through transcriptomic and functional analyses, the following chapter aims to fill in some of the gaps that currently exist in our understanding of the effects of cannabis and the endocannabinoid system on human airway epithelial health; data that is needed to

effectively inform policy, government regulations, and user practices.

3.1 Transcriptomic and barrier responses of human airway epithelial cells exposed to cannabis smoke

Material in this section has been published as part of Aguiar *et al.* (2019).[10] The published manuscript is available here:

> J. A. Aguiar, R. D. Huff, W. Tse, M. R., Stämpfli, B. J. McConkey, A. C. Doxey, and J. A. Hirota. The impact of cigarette smoke exposure, COPD, or asthma status on ABC transporter gene expression in human airway epithelial cells. *Physiological Reports*, 7(20):e14249. 2019.[10] https://doi.org/10.14814/phy2.14249

3.1.1 Introduction

To characterize the human host response to inhaled cannabis smoke, collaborators and I performed a series of *in vitro* functional and transcriptomic experiments with a human airway epithelial cell line (Calu-3) exposed to cannabis smoke, with tobacco smoke used as a positive control. To model cannabis smoke exposure, smoke conditioned media methods that have been validated for tobacco combustion experiments[19,42,45,175–177,188,415] were used. Comparison of differential gene expression patterns from these tobacco smoke conditioned media experiments in Calu-3 cells to bronchial brushings from human tobacco smokers[151,383] and air-liquid interface cultures of primary epithelial cells exposed to mainstream tobacco smoke[152,153,253] revealed overlap, validating the relevance of the model. Using cannabis smoke conditioned media, I observed functional and transcriptional responses that were shared with tobacco smoke. Gene expression pathway analysis revealed that cannabis smoke induced DNA replication and oxidative stress responses. Functionally, cannabis smoke impaired epithelial cell barrier function, antiviral responses, and elevated inflammation. Broadly speaking, the data demonstrate that cannabis smoke exposure is not innocuous and induces transcriptional and functional responses in human airway epithelial cells that may impact lung health similar to tobacco smoke.

3.1.2 Methods

Preparation of cannabis and tobacco smoke extracts

Preparation of smoke extracts performed by Ryan Huff at The University of British Columbia

Cannabis smoke extract (CSE) and tobacco smoke extract (TSE) conditioned media were prepared according to previously published methods.[320] For generation of the TSE, a Kentucky Research Grade Cigarette (Lot: 3R4F – cellulose acetate filter, $\sim$ 0.7 g of dried tobacco leaves) was used. For generation of the CSE, cannabis from Dr. Jonathan Page (University of British Columbia, Vancouver, British Columbia, Canada) (13% THCA strain (w/w), with 0.18% THC, 0.35% THCVA, and 0.18% CBGA, $\sim$ 0.7 g dried cannabis) rolled with card- board filters was used. To prepare the smoke-conditioned media, either 1 cannabis cigarette or 1 tobacco cigarette was smoked into 4 mL of HEPES buffered Eagle's Minimal Essential Medium (EMEM). Smoke extracts were filtered using a 0.22 mL filter. Extracts were standardized by measuring absorbance and diluting with fresh medium to reach a desired dilution (OD260 nm = 0.4045 x dilution factor, 10% dilution = 0.04045 OD260 nm). A single batch of CSE and TSE was generated, aliquoted, and stored at -80°C and used for all subsequent experiments. For RNA-sequencing experiments, 10% CSE and TSE were used. For concentration-response studies, 0.625, 1.25, 2.5, 5, 10, and 20% CSE and TSE were used.

Epithelial cell culture and drugs

Cell culturing performed by Ryan Huff at The University of British Columbia

Calu-3 cells obtained from ATCC (Manassas, Virginia, USA) were grown in EMEM with 10mM HEPES, 10% fetal bovine serum (FBS), and antibiotic-antimycotic, and used between passages 10–20. For exposure experiments, $1x10^6$ or $2x10^5$ Calu-3 cells were seeded onto either 4.7 cm^2 or 0.3 cm^2 polyester Transwell®permeable supports respectively, with a pore size of 0.4 mL. Cells were grown for 20 days to promote cell polarization, with media on both apical and basal sides.[366] FBS was removed from Calu-3 cultures 24 hours prior to the start of the exposures. For exposures, fresh FBS-free EMEM culture media was added to basal chambers and CSE or TSE media diluted with fresh FBS-free EMEM culture media was added to the apical chambers for 24 hours.

Cell viability assay

Cell viability assay performed by Ryan Huff at The University of British Columbia

Cell viability was assessed using a Pierce^{TM} LDH Cytotoxicity Assay Kit (ThermoFisher, Mississauga, Ontario, Canada) according to the manufacturer's instructions.

Barrier function assessment

Barrier function assessment performed by Ryan Huff at The University of British Columbia

Trans-epithelial Electrical Resistance (TEER) was measured using a Millicell ERS-2 Voltohmmeter (EMD Millipore, Etobicoke, Ontario, Canada). Resistance was measured just prior to exposure as well as 24 hours post-exposure and multiplied by the growth area of the inserts (Ωx cm^2).

Cytokine assays

Cell supernatants collected from the apical side of the culture system were analyzed by a 42-plex human cytokine/chemokine protein array (Eve Technologies, Calgary, Alberta, Canada).

RNA-sequencing analysis

Total RNA was extracted using an RNeasy Plus Kit (Qiagen, Toronto, Ontario, Canada). cDNA was prepared at The Centre for Applied Genomics at the Hospital for Sick Children (Toronto, Ontario, Canada). Samples were sequenced on the Illumina HiSeq 2500 instrument with 125 bp paired-end reads to a minimum depth of 30 million reads per sample. Reads were de-multiplexed and trimmed and BCL files generated from the Illumina sequencer were converted to FASTQ files.

After quality control using FastQC (v.0.11.7) and Prinseq (v.0.20.4), sequences were aligned to the human reference genome (hg19) using HISAT2 (v.2.1.0) and assembled into full transcriptomes using StringTie (v.1.3.3b). Samtools (v.1.9) was used to convert and sort HISAT2 output into sorted BAM files for use by StringTie. StringTie was also used to calculate transcript abundances for downstream differential expression analysis using the Ballgown package in R (version 3.4.3) which provides p values, FDR-adjusted p values (q values), and $\log_2$ fold change values for all genes in each comparison. A snakemake-based pipeline called hppRNA (v.1.3.3) was used to combine the above steps into a streamlined work-flow.

Validation of cell culture and exposure methods

To validate the use of Calu-3 cells exposed to smoke extract-conditioned media *in vitro*, four independent tobacco smoke exposure data sets were compared to the results from the TSE versus Control differential expression analysis. Two microarray data sets (GSE4498 and GSE11784)[151,383] and two high-throughput RNA-sequencing data sets (SRP096285 and SRP126155)[152,153] were used for validation. GSE4498 and GSE11784 data sets consist of gene expression data obtained from microarray analysis of mainstream tobacco cigarette smoke-exposed airway epithelial cells compared to cells from healthy individuals (bronchial brushings). SRP096285 and SRP126155 data sets consist of nasal epithelial cells cultured on Transwells and exposed to either 3R4F reference tobacco cigarette smoke extract-conditioned media or air. To determine the correlation between either of these data sets and my own data set, a list of differentially expressed genes from the deposited data set (adjusted $p < 0.05$) was compared to the differentially expressed genes from the TSE versus Control data set (adjusted $p < 0.05$). The intersection of gene names from both lists was determined and these genes were plotted and their correlation was determined using Pearson's r with p value reported. Significance of the gene overlap was determined by a hypergeometric test in R (version 3.4.3).

Functional enrichment and pathway analyses

Lists of significantly differentially expressed genes were identified for CSE versus control and TSE versus control (as determined by Ballgown, FDR q value < 0.05) and genes present in both comparisons were submitted to Enrichr to identify enriched pathways and functional ontologies.[82] Terms were ranked within ontologies by Enrichr's combined score (log(p value) x z-score of the deviation from the expected rank). Expected rank (FDR adjusted p value) was calculated by Enrichr by running the Fisher exact test for random gene sets in order to compute a mean rank and standard deviation from the expected rank for each term in the gene set library.

Statistical analysis

Significant changes in cell viability, TEER, and cytokines were identified through permutation ANOVA followed by Tukey Honest Significant Difference (HSD) *post-hoc* test using the lmPerm package in R (v. 3.4.3). Significant differences between transcriptional expression profiles were identified through analysis of similarities (ANOSIM) using the vegan package in R (v.3.4.3). For all analyses, differences were considered statistically significant when

FDR adjusted p values are less than 0.05, or equivalently with a false-discovery rate of 5%. For all experiments, four independent biological replicates derived from distinct Calu-3 stocks were performed ($n = 4$).

3.1.3 Results

In vitro cannabis smoke exposure-induced changes in gene expression overlap with tobacco smoke

For the described experiments, the transcriptomic changes induced by tobacco smoke were used as positive control stimulus to determine the relative impact of cannabis smoke exposure. Transcriptomic analysis of differentially expressed genes following cannabis or tobacco smoke exposure for 24 hours at 10% dilution revealed a highly significant correlation ($r = 0.695$, $p < 1.0\text{x}10^{15}$) and overlap of shared changes in gene expression (Figure 3.1, $n = 389$, purple triangles).

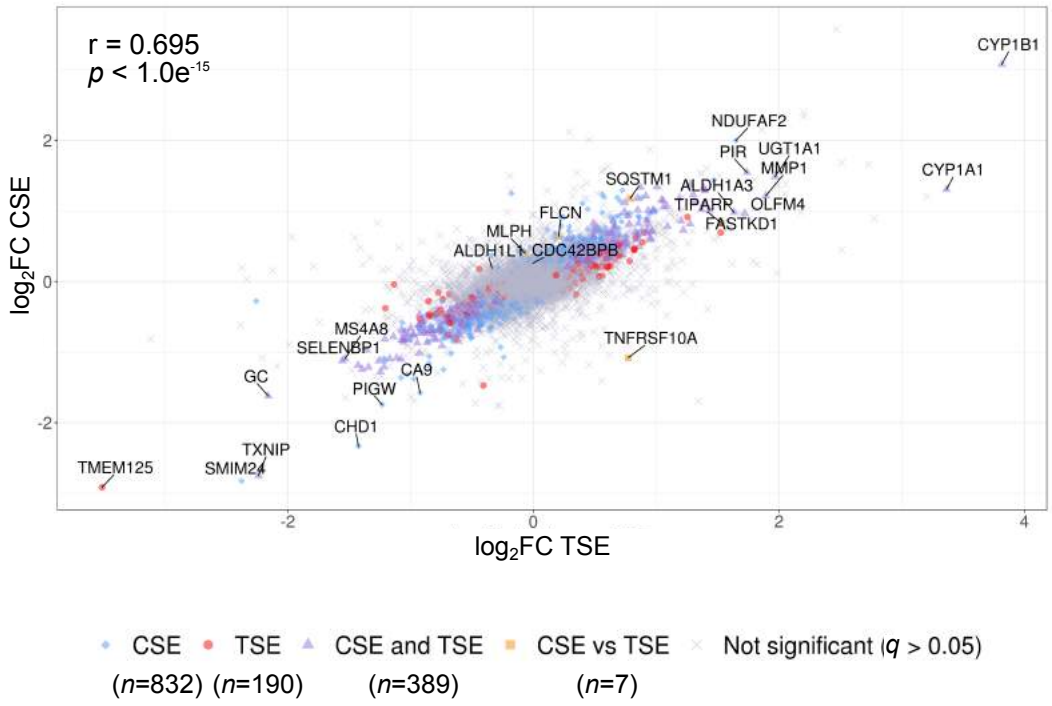


Figure 3.1: Correlation of differential gene expression profiles ($\log_2$ fold change (FC)) versus control, comparing cannabis and tobacco exposures, as determined by Pearson's correlation. Significantly differentially expressed genes in cannabis smoke extract (CSE) only (blue diamonds, $n = 832$), tobacco smoke extract (TSE) only (red dots, $n = 190$), and shared between CSE and TSE (purple triangles, $n = 389$) are highlighted. Only seven genes were identified as significantly differentially expressed between CSE and TSE (orange squares, $n = 7$).

Enriched pathways and functional ontologies were explored based on gene expression patterns using the bioinformatic tool, Enrichr[82] (data not shown). Significantly differentially expressed genes that overlapped between CSE-exposed and TSE-exposed ($n = 389$, purple triangles) cells were analyzed. Analysis of these shared significantly differentially expressed genes revealed increased expression of genes involved in the NRF2 and aryl hydrocarbon receptor pathways that are both involved in responding to oxidative stress and xenobiotic molecules. Moreover, the two genes most up regulated by both smoke exposures were aryl hydrocarbon receptor induced genes *CYP1A1* and *CYP1B1*. I also observed increases in

genes involved with cell replication pathways, cell cycle activating transcription factors E2F1 and MYC, altered barrier function, and lowered antiviral responses in smoke-exposed cells (see Figure 3.3, 3.4).

Minor differences in responses between CSE and TSE reveals that cannabis smoke exposure is able to induce selective changes in gene expression that are not altered with tobacco smoke (Figure 3.1, blue diamonds, $n = 832$) or were different from tobacco smoke (Figure 3.1, orange squares, $n = 7$). Only seven genes (*ALDH1L1*, *CDC42B2B*, *MLPH*, *SQSTM1*, *TIPARP*, *TNFRSF10A*, and *FCLN*) were differentially expressed between CSE and TSE exposure. However, these genes may be of interest in the context of airway epithelial cell biology following smoke exposure. For instance, *TNFRSF10A* (Death Receptor 4, DR4) was significantly down-regulated in CSE relative to TSE exposed cells (4.8 $\log_2$FC). Collectively, these results demonstrate that cannabis smoke exposure impacts transcriptional responses in airway epithelial cells consistent with an oxidative stress phenotype, cell replication, and altered innate immunity.

In vitro tobacco smoke exposure of Calu-3 cells recapitulates differential gene expression patterns observed in human smokers and primary airway epithelial cells

In vitro tobacco smoke exposure models have been extensively used with many experimental variations including cell type (cell line versus primary), culture format (submerged monolayer versus air-liquid interface), or exposure format (smoke conditioned media versus mainstream smoke). Calu-3 cells grown under submerged monolayer conditions were selected as an *in vitro* model for studying the transcriptomic effects of cannabis smoke exposure.[217]

To determine the ability of this model to recapitulate biologically relevant gene expression patterns, I generated an RNA-sequencing transcriptomic data set of Calu-3 cells following 24 hour tobacco smoke exposure (10% TSE) and compared it with previous transcriptomic data sets derived from (1) airway epithelial cells isolated from bronchial brushings of lifetime tobacco smokers; (2) primary human airway epithelial cells grown under air-liquid interface culture conditions and acutely exposed to mainstream tobacco smoke.[151–153,383]

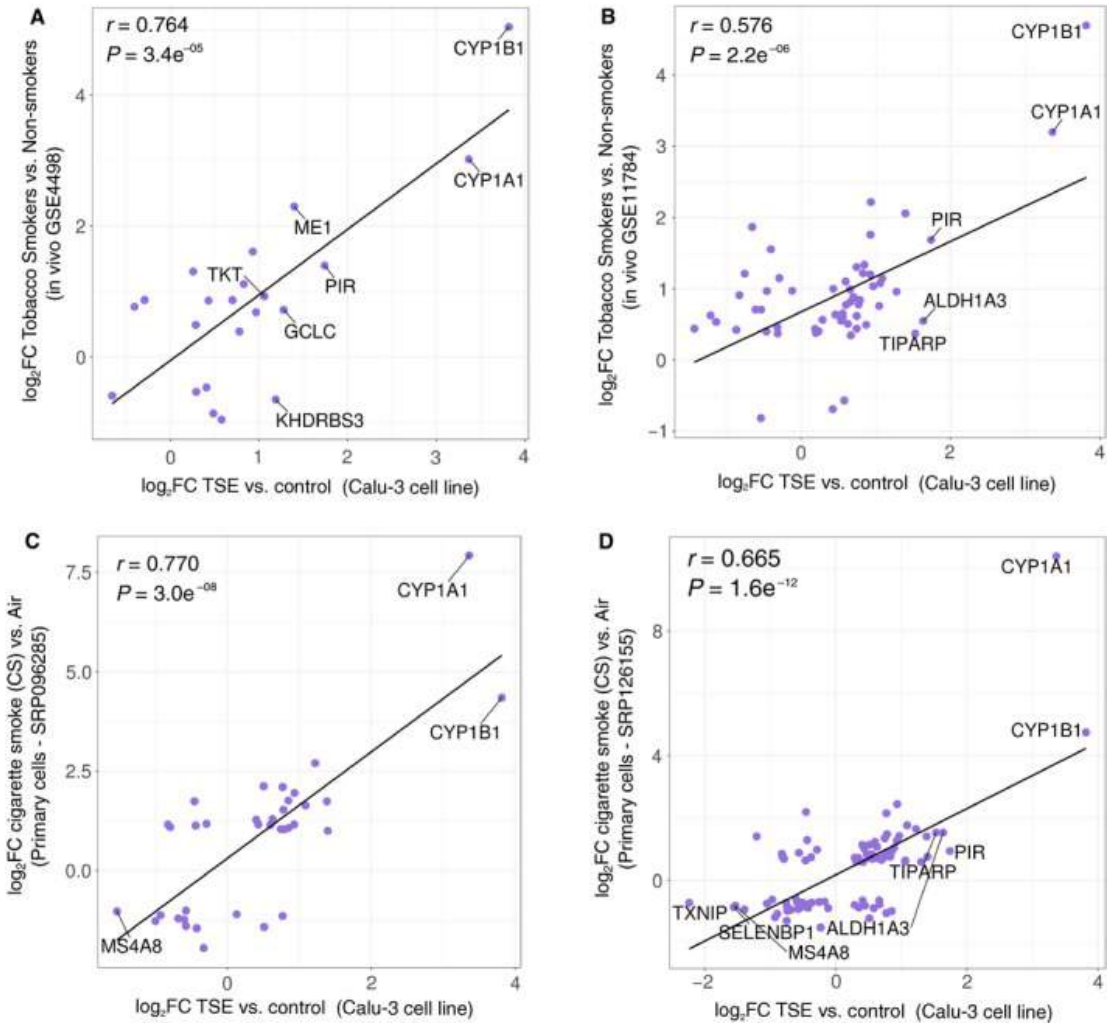


Figure 3.2: Correlation of Calu-3 cell differential gene expression profile ($\log_2$ fold change (FC)) following exposure to tobacco smoke extract (TSE) and differential gene expression profiles between healthy controls and lifetime smokers in the publicly available data sets (A) GSE4498 and (B) GSE11784. Genes that are identified in both data sets and exhibit statistically significant differences in expression are included. Correlations with differential gene expression profiles from primary human airway epithelial cells grown under air-liquid interface culture conditions and exposed to main-stream tobacco smoke using the publicly available data sets (C) SRP096285 and (D) SRP126155.

A correlation ($r = 0.764$, $p = 3.4\text{x}10^{5}$ and $r = 0.576$, $p = 2.2\text{x}10^{6}$) of differentially expressed genes was observed between the transcriptomes of Calu-3 cells exposed to tobacco smoke conditioned media and bronchial brushings obtained by tobacco smokers in two independent data sets (GSE4498 and GSE11784)[151,383] (Figure 3.2 A, B). Furthermore, expression levels of differentially expressed genes were also correlated ($r = 0.770$, $p = 3.0\text{x}10^{8}$ and $r = 0.665$, $p = 1.6\text{x}10^{-12}$) between tobacco smoke exposure Calu-3 cells and air-liquid interface cultures of primary human airway epithelial cells exposed to mainstream tobacco smoke in two independent data sets (SRP096285 and SRP126155)[152,153] (Figure 3.2 C, D). A hypergeometric test shows significant overlap of differentially expressed genes between the transcriptomes of Calu-3 cells exposed to tobacco smoke conditioned media and bronchial brushings from both microarray data sets (GSE4498 and GSE11784, Figure 1 A, B). Significant overlap of differentially expressed genes between the transcriptome of Calu-3 cells exposed to tobacco smoke conditioned media and SRP126155 but not SRP096285 data sets containing air-liquid interface cultures (Figure 1 C, D). Collectively, these results support the validity of using Calu-3 cells under submerged monolayer culture conditions with smoke conditioned media to model and interrogate human airway epithelial cell responses to cannabis.

In vitro cannabis smoke exposure impairs epithelial cell barrier function and related cytokine production without impacting viability

Previous studies have implicated oxidative stress in disrupting epithelial barrier function,[54] therefore my transcriptomic analysis demonstrating oxidative stress suggested that barrier function could be compromised in cannabis smoke exposed cells.

Genes associated with epithelial repair and remodeling were curated from my RNA-seq analysis and plotted in a heat map from the transcriptomic data set (Figure 3.3 A) including *TGFB* (transforming growth factor-β), *TGFBR1* (transforming growth factor-β receptor 1), *EGFR* (epidermal growth factor receptor), *KRT19* (cytokeratin 19), *FN1* (fibronectin), *MMP1* (matrix metalloproteinase 1), and *MMP7* (matrix metalloproteinase 7) (Figure 3.3 A). In parallel, an additional set of genes associated with epithelial cell barrier function were curated including *CDH1* (E-Cadherin), *CTNB1* (β-catenin), *CLDN4* (Claudin 4), *CLDN1* (Claudin-1), and *ZO-1* (Zonus Occludin-1). Cannabis smoke exposure results in both the up-regulation (*FN1*, *MMP1*) and down-regulation (*TGFB1*, *CDH1*, *CTNNB1*) of genes important in epithelial repair, remodeling, and barrier function (adjusted $p < 0.05$). Similar directions of trends were observed for tobacco smoke-induced changes in these genes.

To interrogate transcriptomic changes in the context of cell viability, trans-epithelial

electrical resistance (TEER), and cytokine production at the protein level, a concentration response of cannabis smoke exposure was performed. No changes in cell viability were observed at any concentration of cannabis smoke conditioned media, suggesting cell death was minimally impacted by cannabis smoke (Figure 3.3 B). In the absence of any cell death, cannabis smoke exposure resulted in a concentration-dependent decrease in TEER at 5, 10, and 20% smoke conditioned media (Figure 3.3 C). The viability and barrier function results were comparable with tobacco smoke exposure. The functional consequences of cannabis smoke exposure on cytokine and growth factor protein production important in epithelial cell barrier function were next explored. Cannabis smoke exposure resulted in a concentration dependent increase in TGF-α and PDGF-BB at 20% smoke conditioned media (Figure 3.3 D, E).

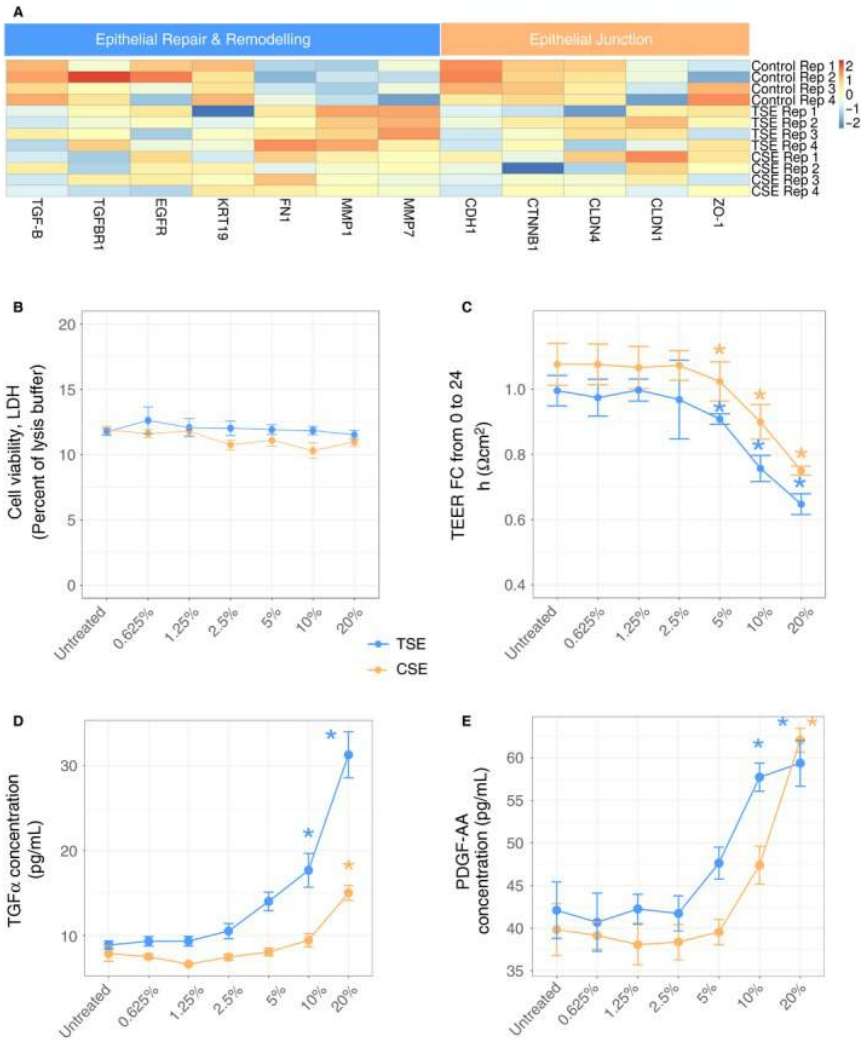


Figure 3.3: Genes involved in (A) airway epithelial repair and remodeling or epithelial junctions were curated from the RNA-sequencing data set performed at 10% CSE or TSE. The expression for each gene is presented for all experimental replicates, with expression for each replicate being scaled by the gene. Calu-3 cells were exposed to increasing concentrations of CSE (orange) or TSE (blue) (control, 0.625%, 1.25%, 2.5%, 5%, 10%, and 20%) for 24 hours with outcome measurements of (B) cell viability assessed by lactate dehydrogenase (LDH) assay, (C) trans-epithelial electrical resistance (TEER) (Ω*cm^2), (D) transforming growth factor-alpha (TGF-α) (pg/mL), and (E) platelet derived growth factor-AA (PDGF-AA) (pg/mL). *: $p < 0.05$ relative to control untreated - Tukey HSD. Error bars represent standard deviation ($n = 4$).

In vitro cannabis smoke exposure attenuates epithelial cell antiviral cytokine responses and induces pro-inflammatory cytokine production

Transcriptomic changes induced by cannabis smoke exposure suggested that antiviral and pro-inflammatory immune responses could be attenuated and induced respectively. Genes associated with antiviral immunity and pro-inflammatory responses were curated from my RNA-seq analysis and plotted in a heat map from the transcriptomic data set (Figure 3.4 A) including *CXCL10*, *CCL5*, interferon stimulatory genes (*IL-6*, *IL-1β*, *IL-8*), and several chemokine receptors (Figure 3.4 A). Cannabis smoke exposure results in both the up-regulation (*IL-8*, *ICAM-1*, *CXCL5*, *CXCL6*) and down-regulation (*CXCL10*, *CCL5*, *RSAD2*, *IFITM3*) of genes important in antiviral immunity and pro-inflammatory responses (adjusted $p < 0.05$). Similar directions of trends were observed for tobacco smoke-induced changes in these curated genes.

Antiviral and pro-inflammatory cytokine transcriptomic changes were analyzed at the protein level in the concentration-response study samples. Cannabis smoke exposure resulted in a concentration dependent trend for a reduction in CXCL10 and a significant reduction in CCL5 at 10 and 20% smoke conditioned media (Figure 3.4 B, C). These trends were comparable with tobacco smoke exposure.

The functional consequences of cannabis smoke exposure on pro-inflammatory cytokine production were explored next. Cannabis smoke exposure resulted in a concentration dependent trend for increased IL-8 and a significant increase in IL-6 at 10 and 20% smoke conditioned media (Figure 3.4 D, E). These trends were also conserved with tobacco smoke exposure.

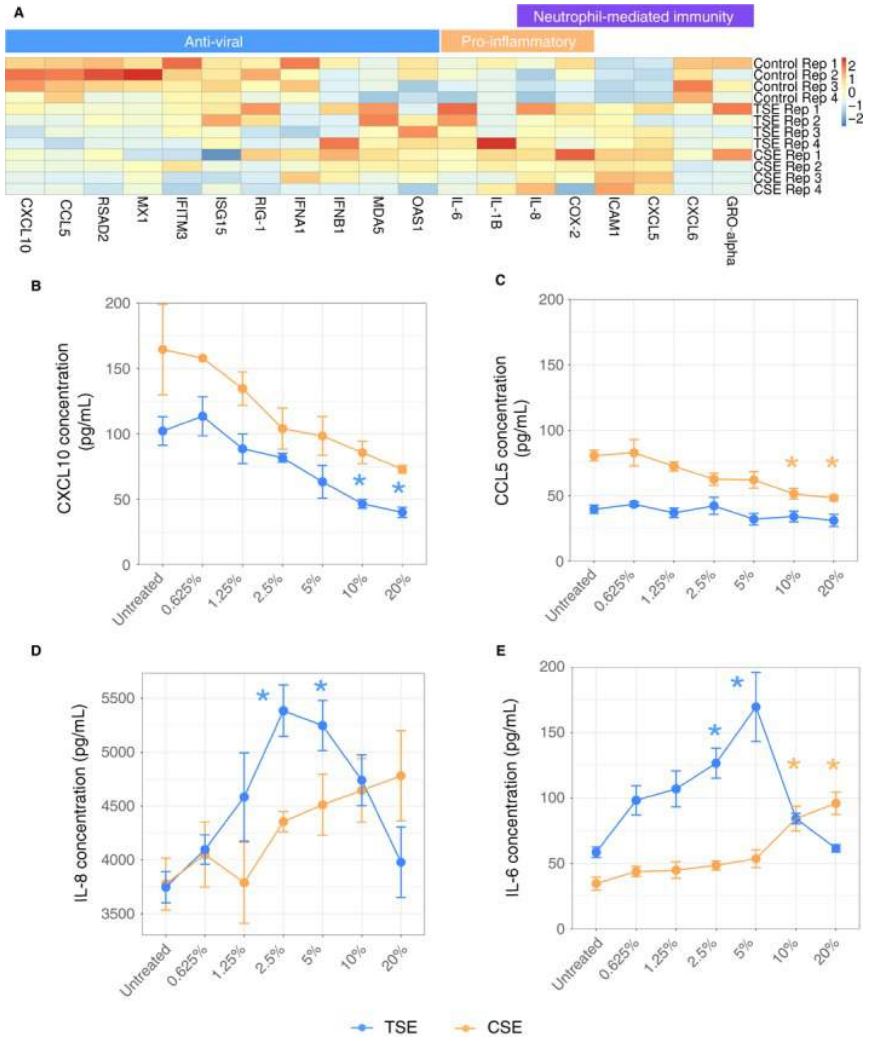


Figure 3.4: Genes involved in (A) airway epithelial anti-viral, pro-inflammatory, or neutrophil mediated immunity were curated from the RNA-sequencing dataset performed at 10% CSE or TSE. The expression for each gene is presented for all experimental replicates, with expression for each replicate being scaled by the gene. Calu-3 cells were exposed to increasing concentrations of CSE (orange) or TSE (blue) (control, 0.625%, 1.25%, 2.5%, 5%, 10%, and 20%) for 24 hours with outcome measurements of (B) interferon gamma induced protein-10 (CXCL10) (pg/mL), (C) regulated on activation, normal T cell expressed and secreted (CCL5) (pg/mL), (D) IL-8 (pg/mL), and (E) IL-6 (pg/mL). *: $p < 0.05$ relative to control untreated - Tukey HSD. Error bars represent standard deviation ($n = 4$).

3.1.4 Discussion

Global cannabis use is an important public health issue that would benefit from experimental evidence to inform policy, government regulations, and user practices. Analyses that interrogate cannabis in parallel with a positive control of tobacco smoke will help provide clinically relevant evidence and context.[27,193,374,381,420] The *in vitro* functional and transcriptomic experiments with Calu-3 human airway epithelial cells exposed to cannabis smoke demonstrate changes in gene expression signatures related to DNA replication and oxidative stress responses, an impairment of barrier function and antiviral immune responses, and an augmented pro-inflammatory cytokine profile. Importantly, all cannabis-induced responses were observed in tandem with the positive control of tobacco smoke, suggesting potential parallel implications for lung health.

In vitro experiments using human airway epithelial cell lines and primary cell samples under different growth conditions have been essential methods in tobacco smoke exposure science,[42,45,152,153,175–177,188,253,415] while no such diversity and baseline of data currently exists for cannabis smoke exposure science. Potential exists for differences in experiment design (e.g. submerged monolayer cell line exposed to smoke-conditioned media versus air-liquid interface primary airway epithelial cell exposed to mainstream smoke), which may confound and limit translation of experimental results for any smoke exposure. To assess the validity of this *in vitro* model of Calu-3 cells exposed to smoke conditioned media, I compared the differentially expressed genes from my tobacco exposure experiments with publicly available data sets from gold standard bronchial brushings in addition to an air-liquid interface model of primary human airway epithelial cell cultures. Differential gene expression analyses demonstrate significant correlation between the proposed *in vitro* model of Calu-3 cells with *in situ* bronchial brushings and *in vitro* air-liquid interface cultures of primary human airway epithelial cells. Importantly, significant correlation existed between the study model and *in situ* profiles and more complex *in vitro* models, despite the use of the Calu-3 cell line (relative to primary cells) and smoke-conditioned media (relative to main- stream smoke). The results from these cannabis smoke exposure studies are therefore potentially reflective of the *in situ* condition, although this should be validated with clinical studies.

The function of the airway epithelium is to provide the lung a physical and immunological barrier to the environment.[179,288] Any perturbation in the airway epithelium may lead to host susceptibility to infection and lung pathology or disease development. The data in this study demonstrate that cannabis smoke is able to induce mild impacts on barrier function, measured by TEER, without impacting cell viability. The mechanism(s) by which TEER is reduced by cannabis smoke were not determined in this study, but in the absence of

changes in cell viability it is possible that cell-cell junctions could have been disrupted as has been reported for tobacco smoke,[334] possibly by oxidative stress.[54] The down-regulation of E-cadherin (CDH1) by cannabis smoke has been previously observed in lung cells exposed to tobacco smoke and associated with the induction of epithelial-to-mesenchymal transition, migration, and invasion phenotypes[270] that could be important in the context of lung health of the cannabis smoker. Disruption of E-cadherin may contribute to β-catenin shuttling from membrane junction sites to transcriptional locations in the nucleus to facilitate gene expression associated with repair.[265] Importantly, aberrant β-catenin signaling in cells is associated with oncogenic gene expression signatures.[112] Independent of any aberrant transcriptional regulation resulting from disrupted epithelial barrier function; a reduced impedance of the physical barrier of the airway epithelium offers easier access to the lung for opportunistic insults from the outside world.

To complement the mechanical barrier of the lung, the airway epithelium is able to produce and host-defence peptides and antiviral mediators to protect from pathogens.[179,288] Tobacco smoke has been reported to compromise the ability of airway epithelial cells to effectively control both bacterial and viral insults.[19,111,177,320] The data in this study demonstrate striking similarities in the epithelial immune profile in response to cannabis and tobacco, suggests that the former will impact host defences. Tobacco has been demonstrated to impact host defence peptide induction in airway epithelial cells by non-typeable *Haemophilus influenzae* with a concomitant increase in IL-8 expression.[20] The experimental data set in this study did not include a pathogen challenge, precluding my ability to confirm the tobacco smoke induced suppression of host defence peptides or to extend the results to cannabis smoke. In the context of viral exposures, type I interferons (IFNs) are capable of rapidly inducing interferon stimulated genes (ISGs) through the type I IFN receptor to help tackle various components of virus replication, assembly and budding.[138] Tobacco smoke has been shown to impact antiviral immunity in airway epithelial cells in response to human rhinovirus-16, with a reduction in CXCL10 and CCL5.[111] Tobacco smoke-induced reduction in CXCL10 and CCL5 is associated with greater rhinovirus production. Tobacco smoke has also been demonstrated to impair airway epithelial cell antiviral immunity mediated by IFN-γ in response to respiratory syncytial virus exposure,[264] which in turn could impact CXCL10 and CCL5 production.[293] The cytokine and transcriptomic data in this study confirm the tobacco smoke impairment of CXCL10 and CCL5 and antiviral ontologies and show that this is conserved for cannabis smoke. It is further demonstrated using my previous transcriptomics data set that a diverse selection of ISGs were also attenuated with both cannabis and tobacco, consolidating a common phenotype between both smoke exposures. Future co-culture experiments with pathogen challenges would help reveal the complex interaction between cannabis and tobacco

smoke induced immune responses in the epithelium and down-stream immune cell phenotype and function. Collectively, although the experimental design lacks the mechanistic linkage between cannabis exposure and increased susceptibility to viral or bacterial infections, the data strongly mirror those for tobacco, which has been mechanistically linked to compromised host immunity to pathogens.

Extensive evidence exists that tobacco and biomass smoke exposure, the latter generated from dried wood, animal dung, or charcoal, are risk factors for the development of chronic bronchitis, emphysema, and lung cancers.[137,179,205,223] In contrast, the existing evidence suggests that repeated cannabis smoke exposure results in a chronic bronchitis phenotype with little evidence of emphysema.[17,371,374,420] Furthermore, unlike tobacco and biomass exposure, which are accompanied by a dose-dependent risk for development of lung cancers, a similar relationship has not been observed for repeat cannabis users despite the presence of carcinogens in cannabis smoke.[434] Of particular note, it was observed that cannabis smoke significantly up-regulated the proto-oncogene *MYC*, which has previously been observed in airway epithelial cells exposed to tobacco smoke[244] or benzo[a]pyrene, a known carcinogen that has been identified in both tobacco and cannabis smokes.[115,266] Despite the potential for cannabis smoke to up-regulate proto-oncogenes, why cannabis smoke exposure is only tenuously linked to cancer[373] remains to be determined and should be explored in additional cohorts in jurisdictions where cannabis has been legalized.

Tobacco smoke exposure experiments have used standardized research source material to ensure experimental reproducibility and robustness. In contrast, cannabis experiments have not benefited from a widely accessible and chemically defined source material. For this reason, a cannabis strain that was representative of that available in the medicinal cannabis market in Canada, that included 13% THCA (w/w), 0.18% THC, 0.35% THCVA, and 0.18% CBGA with no levels of CBD, was used. These results must therefore be interpreted based on this initial chemical composition and care should be taken to generalize that all cannabis strains will induce similar responses. Indeed, increasing evidence suggests that there may be complex interactions between THC and CBD via the CB1 and CB2 cannabinoid receptors that could impact the immunomodulatory functions of cannabis smoke.[57,283,360] Exposure studies could collectively benefit by reporting the chemical composition of the strain that was used to help facilitate interpretation of data generated.

Methods incorporating the use of smoke conditioned media have been used extensively for tobacco research.[19,42,45,175–177,188,415] Smoke conditioned media methods typically filter coarse particulates and extract water-soluble components of smoke combustion. The reduction in compositional complexity of smoke conditioned media relative to mainstream smoke may be important, although the data in this study and those of others[152,153,253] demonstrate that major transcriptional changes are conserved in either model system

(smoke-conditioned media or mainstream smoke) and both are reflective of *in situ* human biology.[151,383] Importantly, the presence or absence of filters in tobacco (cellulose acetate) and cannabis smoke (cardboard) exposure models should be considered in interpreting the present data and designing future experiments.[28] Collectively, the data validate smoke conditioned media models and suggest that they will provide insight into the impacts of cannabis smoke exposure on airway epithelial cell biology.

In conclusion, the data in this study demonstrate striking similarities in the impacts of cannabis and tobacco smoke on airway epithelial cell barrier function, cytokine profile, and gene expression signatures. Despite the arrival of cannabis legalization, the data suggest that cannabis smoke exposure still poses a significant health risk and warrants ongoing study to build a body of evidence to support public policy, government regulations, and user practices.

3.2 Effect of long-acting β-agonist/glucocorticoids on human airway epithelial cell cytokine, transcriptomic, and oxidative stress responses to cannabis smoke

Material in this section has been published as part of Huff *et al.* (2020).[178] The published manuscript is available here:

> R. D. Huff, J. A. Aguiar, W. Tse, M. R., Stämpfli, B. J. McConkey, A. C. Doxey, and J. A. Hirota. Effect of long-acting β-agonists/glucocorticoids on human airway epithelial cell cytokine, transcriptomic and oxidative stress responses to cannabis smoke. *European Respiratory Journal Open*, 6(1):00265. 2020.[178] https://doi.org/10.1183/23120541.00265-2019

3.2.1 Introduction

To manage chronic lung diseases, an individual may be prescribed a broad, anti-inflammatory long-acting β-agonist (LABA)/glucocorticoid (GC) combination treatment.[154,292] LABA/GCs are transcriptionally active in airway epithelial cells resulting in broad anti-inflammatory activities that are believed to be important in controlling infections and chronic lung disease

daily management and to prevent exacerbations.[126,127,203] Importantly, the efficacy of LABA/GC therapies may be compromised by tobacco smoke exposure.[154,320] It remains to be determined whether cannabis smoke exposure similarly compromises the efficacy of LABA/GC anti-inflammatory activities. As cannabis use becomes more universally accepted, it is important to understand the effects that cannabis use has on airway health and the interventions used to manage lung inflammation and disease. Given the similarities between tobacco and cannabis smoke,[10] I examined whether cannabis smoke exposure similarly attenuates LABA/GC transcriptomic responses and inflammatory mediator release in human airway epithelial cells.

3.2.2 Methods

Smoke extract preparation, cell culturing, oxidative stress assessment performed by Ryan Huff at The University of British Columbia

Cannabis smoke extract (CSE)- and tobacco smoke extract (TSE)-conditioned media were prepared according to previously published methods.[320] TSE was generated from Kentucky Research Grade Cigarettes with an intact filter (lot 3R4F), and CSE from cannabis rolled with cardboard filters and sourced from Jonathan Page (University of British Columbia, Vancouver, Canada) (13% (w/w) tetrahydrocannabinolic acid strain with 0.18% THC, 0.35% tetrahydrocannabivarinic acid and 0.18% cannabigerolic acid; ∼ 0.7 g dried cannabis). Extracts were prepared by bubbling cannabis or tobacco smoke through 4 mL HEPES-buffered Eagle's minimal essential medium, filtering (0.22 µm) and diluting with fresh medium (10% dilution, optical density at 260 nm 0.04045). Calu-3 lung epithelial cells on Transwells®(Corning Inc., Corning, NY, USA) were apically exposed to 10% of either CSE or TSE with and without 10 nM of the LABA formoterol and 100 nM of the GC budesonide (Form/Bud) for 24 hours. Total RNA was extracted and sent to The Centre for Applied Genomics at the Hospital for Sick Children (Toronto, Canada) for RNA sequencing. Apical cell supernatants were assessed using a multiplex cytokine/chemokine protein array by Eve Technologies (Calgary, Canada). Oxidative stress was assessed using 2',7'-dichlorodihydrofluorescein diacetate (H_2DCFDA) and CellROX Green (Thermo Fisher Scientific, Waltham, MA, USA), counter-stained using Hoechst 3342 and quantified using a plate reader. All experiments had four biological replicates. Significant differences in cytokines and oxidative stress were identified through permutation ANOVA followed by a Tukey HSD *post-hoc* test using the lmPerm package in R (version 3.4.3). Significant differences between transcriptional expression profiles were identified through analysis of similarities using the vegan package in R (version 3.4.3).

3.2.3 Results

Exposure to 10% CSE or TSE and/or Form/Bud for 24 hours did not significantly change cell viability as assessed by lactate dehydrogenase release assays (not shown). Like TSE, CSE significantly elevated pro-inflammatory cytokine interleukin-8 (IL-8)/CXCL8 and epithelial cell repair mediator transforming growth factor-α (Figure 3.5 A, B). Intervention with Form/Bud prevented CSE and TSE increases in IL-8/CXCL8 release and, to a lesser extent, TGF-α. A significant decrease in interferon-γ-induced protein 10 (IP-10)/CXCL10 (Figure 3.5 C) in response to CSE and TSE exposure was also observed, a decrease that was augmented in response to Form/Bud intervention.

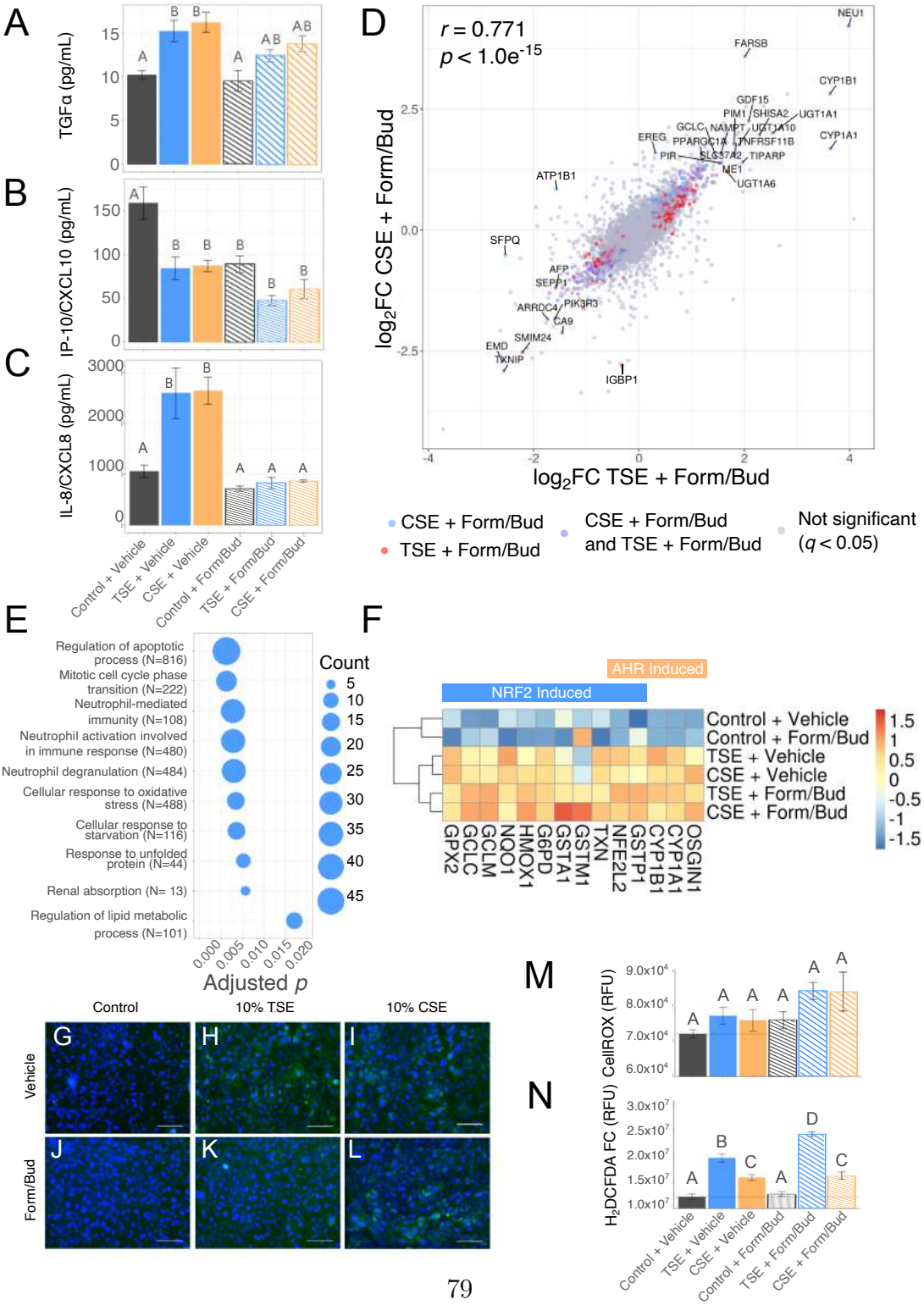

A
TGFα (pg/mL)
B
IP-10/CXCL10 (pg/mL)
C
IL-8/CXCL8 (pg/mL)
Control + Vehicle
TSE + Vehicle
CSE + Vehicle
Control + Form/Bud
TSE + Form/Bud
CSE + Form/Bud
D
r = 0.771
p < 1.0e-15
log2FC CSE + Form/Bud
log2FC TSE + Form/Bud
CSE + Form/Bud
TSE + Form/Bud
CSE + Form/Bud and TSE + Form/Bud
Not significant (q < 0.05)
E
Regulation of apoptotic process (N=816)
Mitotic cell cycle phase transition (N=222)
Neutrophil-mediated immunity (N=108)
Neutrophil activation involved in immune response (N=480)
Neutrophil degranulation (N=484)
Cellular response to oxidative stress (N=488)
Cellular response to starvation (N=116)
Response to unfolded protein (N=44)
Renal absorption (N= 13)
Regulation of lipid metabolic process (N=101)
Count
Adjusted p
F
AHR Induced
NRF2 Induced
Control + Vehicle
Control + Form/Bud
TSE + Vehicle
CSE + Vehicle
TSE + Form/Bud
CSE + Form/Bud
GPX2
GCLC
GCLM
NQO1
HMOX1
G6PD
GSTA1
GSTM1
TXN
NFE2L2
GSTP1
CYP1B1
CYP1A1
OSGIN1
Control
10% TSE
10% CSE
Vehicle
Form/Bud
G
H
I
J
K
L
M
CellROX (RFU)
N
H2DCFDA FC (RFU)

Figure 3.5: Calu-3 cells were exposed to either control, 10% CSE (orange) or 10% tobacco smoke extract (TSE) (blue) with either vehicle or Form/Bud for 24 hours. To assess the impact of Form/Bud on inflammatory mediators, cytokine release of (A) TGF-α, (B) IP-10/CXCL10 and (C) IL-8/CXCL8 was measured. (D) The effect of Form/Bud intervention and CSE or TSE exposure on airway epithelial gene expression was assessed by RNA sequencing. The $\log_2$ fold change (FC) of CSE+Form/Bud versus vehicle+Form/Bud was plotted against the $\log_2$FC of TSE+Form/Bud versus vehicle+Form/Bud to reveal significant TSE- and CSE-specific genes under Form/Bud intervention conditions (r = 0.771; $p < 1\times$10-15). Genes that were significantly differentially expressed ($q < 0.05$) in CSE+Form/Bud and TSE+Form/Bud versus vehicle+Form/Bud are highlighted in purple, where as CSE- and TSE-specific genes are highlighted in blue and red respectively. To examine the similarity of CSE+Form/Bud and TSE+Form/Bud transcriptomic profiles, correlation of differential gene expression was determined by Pearson's correlation. (E) Functional enrichment analysis of genes shown to be significantly differentially expressed ($q < 0.05$) in CSE+Form/Bud and TSE+Form/Bud versus vehicle+Form/Bud. Top 10 gene ontology biological processes were ranked by decreasing adjusted p value, with number of significantly differentially expressed genes (count) contributing to the total number of genes associated with the given pathway (n) denoted by the size of circle. (F) Expression of genes involved in NRF2 and AHR responses. (G-L) Oxidative stress was assessed using CellROX and m) the fluorescent intensity was quantified n) with complementary quantification using 2',7'-dichlorodihydrofluorescein diacetate (H_2DCFDA). All experiments were performed using four biological replicates from serial passages. Different letters indicate significantly different group means (Tukey Honest Significant Difference, $p < 0.05$). There are no significant differences between means sharing a common letter. Error bars represent standard error of the mean (SEM). Scale bars = 100 µm. RFU: relative fluorescence unit.

Next, RNA sequencing data was interrogated to address the efficacy of LABA/GC intervention on attenuating CSE- and TSE-induced transcriptomic responses observed in my previous study.[10] First, I observed that the most up-regulated gene exclusive to Form/Bud treatment was *HSD11B2* (corticosteroid 11-β-dehydrogenase) (vehicle + Form/Bud $\log_2$FC 34.80, CSE + Form/Bud $\log_2$FC 33.99 and TSE + Form/Bud $\log_2$FC 27.32), indicating that intervention was efficacious on modulating glucocorticoid signalling (data not shown). Second, to identify expression patterns unique to each smoke exposure + LABA/GC intervention, I directly compared these data sets to identify differentially expressed genes and visualized their $\log_2$ fold changes relative to untreated + Form/Bud (Figure 3.5 D). Form/Bud intervention in smoke-exposed cells induced a highly correlated transcriptomic response (r = 0.771, $p < 1.0$x10^{-15}), with 801 differentially expressed genes in CSE +

Form/Bud versus untreated + Form/Bud and 1105 differentially expressed genes in TSE + Form/Bud versus untreated + Form/Bud. Third, *NEU1* was the most significantly increased gene in both CSE + Form/Bud and TSE + Form/Bud compared to vehicle + Form/Bud exposed cells, and aryl hydrocarbon receptor (AHR) genes *CYP1A1* and *CYP1B1* remained up-regulated in both CSE + Form/Bud- and TSE + Form/Bud-exposed cells.

To further interrogate genes significantly differentially expressed in both CSE + Form/Bud- and TSE + Form/ Bud-exposed cells, I performed a functional enrichment analysis (Figure 3.5 E). Interestingly, regulation of apoptotic process (rank 1, 45 out of 816 genes, gene ontology (GO:0042981), neutrophil-mediated immunity (rank 3, 30 out of 108 genes, GO:0002446), and cellular response to oxidative stress (rank 6, 13 out of 488 genes, GO:0034599) were included in the top 10 GO terms ranked by adjusted p value. To functionally characterise the "cellular response to oxidative stress" term, collaborators performed additional *in vitro* experiments and analysis of oxidative stress-related gene expression data[179,354] (Figure 3.5 F). Interestingly, several curated genes (e.g. *GCLC* and *GCLM*) were increased beyond CSE or TSE exposure with Form/Bud intervention. Furthermore, when overall oxidative stress was examined using CellROX (Figure 3.5 G, M) and H_2DCFDA (Figure 3.5 N) *in vitro* imaging assays, reactive oxygen species (ROS) generation was observed in both CSE + Form/Bud- or TSE + Form/Bud-exposed cells.

3.2.4 Discussion

Tobacco smoke has long been associated with increased symptoms and severity of chronic lung diseases that include asthma, pulmonary fibrosis and COPD, and attenuated GC responses in human lung epithelial cells.[320] Chronic bronchitis and airway remodelling/scarring due to persistent inflammation and oxidative stress from recurring exposure to tobacco smoke are known to be important pathological mechanisms in the progression of COPD and asthma.[154,170] It is demonstrated here for the first time that a Form/Bud intervention of human lung cells exposed to cannabis smoke suppressed increased inflammation (IL-8/CXCL8) and epithelial repair mediators (TGF-α), while expression of oxidative stress genes remained elevated. Although Form/Bud intervention reduced the expression of IL-8/CXCL8, CSE and TSE exposure may still be modulating epithelial immune responses, as seen by the continued enrichment of neutrophil-related GO terms. Furthermore, the persistence of increased generation of ROS and expression of the NRF2 oxidative stress response genes in the presence of LABA/GC medications used in chronic respiratory disease management may further impact long-term disease development and management.

This study has several limitations in the experimental design. The model used is a single exposure to smoke-conditioned media and may not accurately represent chronic smoke exposure conditions, which would require repeated exposures over time *in vitro*. Despite this limitation, I emphasise that my previous publication using this model demonstrated strong significant correlations of gene expression with *in situ* bronchial brushes from human smokers and primary air–liquid interface cultures exposed to tobacco smoke.[10] Additional experiments using primary human airway epithelial cells grown under air–liquid interface culture conditions and exposed to whole-smoke extract should be explored to complement the present data. I recognise that there are substantial differences in cannabis strains available on the medicinal and recreational market, and therefore chose a representative hybrid strain available on the medicinal cannabis market in Canada. Despite these limitations, the robust changes observed in this data suggest that cannabis smoke exposure still poses a significant health risk, and warrants ongoing study to build a body of evidence to support public policy, government regulations and individual user practices.

3.3 Expression of the endocannabinoid system in the human airway epithelial cells: impacts of sex and chronic respiratory disease status

Material in this section has been published as part of Fantauzzi *et al.* (2020).[113] The published manuscript is available here:

> M. F. Fantauzzi, J. A. Aguiar, B. J.-M. Tremblay, M. J. Mansfield, T. Yanagihara, A. Chandiramohan, S. Revill, M. H. Ryu, C. Carlsten, K. Ask, M. R., Stämpfli, A. C. Doxey, and J. A. Hirota. Expression of endocannabinoid system components in human airway epithelial cells: impact of sex and chronic respiratory disease status. *European Respiratory Journal Open*, 6(4):00128. 2020.[113] https://doi.org/10.1183/23120541.00128-2020

3.3.1 Introduction

The mechanism(s) responsible for the clinical observations resulting from cannabis inhalation, whether beneficial acute or detrimental chronic exposures, remain elusive, but are likely to be influenced by the endocannabinoid system.[66,140,386]

The endocannabinoid system is responsible for mediating the pharmacological effects of the endogenous cannabinoids anandamide and 2-arachidonylglycerol (2-AG) and phytocannabinoids present in cannabis.[102,300,365] The first identified cannabinoid receptors were CB1 and CB2, both G-protein coupled receptors that modulate downstream cyclic-AMP signaling by inhibiting adenylyl cyclase activity.[254,268] The mechanism of acute phytocannabinoid-induced bronchodilation is suggested to be mediated via CB1 receptors regulating neural control of airway tone,[66,140] which may have benefits for management of bronchoconstriction in asthma or COPD. In addition to CB1 and CB2, the transient receptor potential vanilloid-1 (TRPV1) and the de-orphanized receptor GPR55 have been identified as receptors for cannabinoids.[329,446] Although the endocannabinoid signaling pathway is present throughout multiple tissues and organ systems, the lungs are of particular interest as they are the organ system targeted by cannabis inhalation. Airway epithelial cells play a critical role in lung health by acting as the first line of defence against pathogens and inhaled insults.[159,163,288] Airway epithelial cells carry out a number of functions such as providing a physical barrier against microbial infiltration, maintaining the inflammatory microenvironment, and releasing immune mediators to recruit leukocytes to the site of insult. It has been demonstrated that the inhalation of air pollution, tobacco smoke, and cannabis smoke can compromise airway epithelial function.[10,178,179] Notably, recent findings show that cannabis smoke exposure can lead to impaired airway epithelium barrier integrity, attenuated antiviral capacity, and exacerbated inflammatory response to immune challenges.[10,178] However, the contribution of the endocannabinoid system to these observations has not been defined. The primary cannabinoid receptors, CB1 and CB2, have been shown to be expressed in the respiratory mucosa[93,121] and human airway epithelial cells are responsive to THC and anandamide.[342,343] Additional components of the endocannabinoid system including MAPK, PI3K, and protein kinase A signaling pathways downstream of receptors and enzymes responsible for cannabinoid metabolism have not been explored in an integrated fashion in human airway epithelial cells. An examination of the entire endocannabinoid system in human airway epithelial cells is required to better understand which components are dominant and likely to be functionally relevant in response to inhaled cannabis smoke. Furthermore, it remains possible that sex and disease status impact the endocannabinoid system expression, which may have functional consequences in distinct populations of cannabis consumers. Therefore, a generalized understanding of how a cannabis consumer responds to smoke exposure is likely to be insufficient and should include both females and males in analyses and consider both healthy individuals and those that may have underlying medical needs.

To begin my interrogation of the endocannabinoid system in human airway epithelial cells, a 32-gene endocannabinoid signature encompassing ligand recognition, signaling,

and metabolism was generated. I set out to examine if the expression of this 32-gene endocannabinoid signature was present in human airway epithelial cells and whether this was impacted by sex or disease status. The importance on examining sex and disease status on the endocannabinoid system is due to the possibility that specific populations may experience differential effects of cannabis, either positive or negative. To complete this study I used a bioinformatic approach to analyze gene expression in 1090 unique human subject samples of airway epithelial cells isolated via bronchial brushing that included samples from males and females and individuals with asthma or COPD. My bioinformatic approach is complemented with validation and confirmation of CB1, CB2, and TRPV1 in human airway epithelial cells at the protein level *in situ* and *in vitro*. Lastly, my bioinformatic observation that TRPV1 gene expression is elevated in airway epithelial cells isolated from asthmatics is validated by performing confirmatory immunoblot analysis on primary human airway epithelial cells. Collectively, these results demonstrate that an intact endocannabinoid system is expressed in human airway epithelial cells and that both sex and disease status impact expression, which may have functional consequences that lead to differential responses in distinct populations of cannabis consumers.

3.3.2 Methods

Human ethics

All studies using primary human lung material and blood were approved by Hamilton Integrated Research Ethics Board or University of British Columbia Human Research Ethics Board.

Primary human airway epithelial cells

Cell preparation performed by Matthew Fantauzzi at McMaster University

Primary human airway epithelial cells isolated via bronchial brushings from consented healthy or asthmatic subjects were grown in PneumaCult ExPlus (Stemcell Technologies, Vancouver Canada) under submerged monolayer culture conditions and used in between passage 1 and 4. Where relevant, asthma diagnosis was confirmed with methacholine challenge and PC20 analysis as per ATS guidelines.

Human whole lung tissue

Cell preparation performed by Matthew Fantauzzi at McMaster University

Non-involved tissues from lung cancer cases were used. Lungs were homogenized using a mechanical homogenizer (Omni International, Waterbuy CT), lysed in 1X lysis buffer supplemented with complete protease inhibitors (Roche), and the supernatant was collected for immunoblots.

Human Peripheral Blood Mononuclear Cells (PBMC)

Cell preparation performed by Matthew Fantauzzi at McMaster University

Human PBMC were isolated from the peripheral blood of healthy volunteers using Ficoll (Sigma-Aldrich) density centrifugation in Greiner LeucoSep-tubes (Sigma) according to the manufacturer's recommendations.

Immunohistochemistry

Immunohistochemistry performed by Matthew Fantauzzi at McMaster University

Formalin fixed paraffin embedded human lung tissue from healthy subjects were used for localization of CB1, CB2, and TRPV1/VR1. Three-micron thick sections were cut and stained for CB1 (Abcam - Ab23703 at 1:1000), CB2 (Abcam – Ab3561 at 1:50), and TRPV1/VR1 (Abcam – Ab3487 at 5 µg/mL). All staining was performed on a Leica Bond RX system with Leica Bond reagents and heat-induced antigen retrieval in citrate buffer at pH 6. Digital slide scanning was performed using an Olympus VS120-L100 Virtual Slide System at 40X magnification with VS-ASW-L100 V2.9 software and a VC50 colour camera.

Immunoblots

Immunoblots performed by Matthew Fantauzzi at McMaster University

Immunoblots confirming antibody staining and protein expression in human airway epithelial cells were performed using Biorad stain free 4-20% pre-cast gradient gels and imaged on a Biorad ChemiDoc XRS+ Imaging system. For each immunoblot, 40 µg of protein was added per lane. CB1 (Abcam - Ab23703 at 1:500), CB2 (Abcam – Ab3561 at 1:100), and TRPV1 (Abcam – Ab3487 at 1:500) were diluted in 5% skim milk/Tris buffered saline with 0.1% Tween-20. Primary antibody detection was performed using an anti-rabbit-HRP conjugated secondary (Cell Signaling Technology, #7074) at 1:1000 for 50 minutes at room temperature. Visualization was performed using Clarity™ Western ECL Substrate (Bio-Rad) (CB1 and TRPV1) and SuperSignal™ West Femto Maximum

Sensitivity Substrate (ThermoFisher) (CB2). Total protein loading images were collected as a confirmation of equal protein loading between sample types.[144]

Gene expression data set curation, normalization, and statistical analyses

Candidate genes to be analyzed were selected based on known relevance in the endocannabinoid signaling pathway. Additional genes and proteins implicated to be involved in cannabis-associated disorders and novel cannabinoid based therapeutic approaches were included and described in annotated table format (Table 8).[9,53,122,155,199,252,254,268,329,410,446,58,59,77,89,104,172,200,206,210,278,307,328]

Public microarray experiments using Affymetrix chips (HG-U133 Plus 2, HuEx-1.0-st-v1 and HuGene-1.0-st-v1) on airway epithelial cell samples from healthy individuals or those with asthma or COPD were selected from the NCBI Gene Expression Omnibus (GEO) database. Healthy samples were further filtered by removing former or current smokers. This resulted in a total of 1090 individual samples from 27 experiments (Table 9) that included samples from 616 healthy subjects, 136 subjects with asthma, and 338 subjects with COPD.[21,51,61,62,65,73–75,87,123,151,157,276,312,335,346,368,382,383,387,402,403,405,406,416,418,428,439]

Within each sample population, sex was reported for a subset of samples (Healthy: 103 females/227 males, Asthma: 34 females/28 males, COPD: 48 females/93 males).

For all data set samples, raw intensity values and annotation data were downloaded with the R statistical language (version 3.6.1) using the GEOquery R package (version 2.52.0).[99] Probe definition files were downloaded from the Brainarray database (version 24).[97] To obtain processed microarray gene expression values unaffected by probe CG compositional biases, the Single Channel Array Normalization (SCAN) method was used via the SCAN.UPC R package (version 2.26.0)[301] using annotation data from the Bioconductor project (version 3.9).[174] All $\log_2$-transformed gene expression data were unified into a single data set, and only genes detected in all three platforms ($n = 16543$) were kept for subsequent analyses. Correction of experiment-specific batch effects was performed using the ComBat method[195] implemented in the sva R package (version 3.32.1),[229] with disease status and sex supplied as covariates. Principal component analysis (PCA) was performed using the probabilistic PCA method in the pcaMethods R package (version 1.76.0).[361] Gene expression levels were tested for significant differences via Student's T-test with a Benjamini-Hochberg multiple testing correction using the stats R package (version 3.6.1). Gene expression box plots, heat maps, and PCA plots were generated with the ggplot2 R package (version 3.2.1).

3.3.3 Results

Human airway epithelial cells express CB1, CB2, and TRPV1, *in situ* and *in vitro*

A curated list of genes involved in cannabinoid signaling was generated (herein called the 32-gene endocannabinoid signature) to provide a focused overview of this pathway in human airway epithelial cells (Figure 3.6 and Table 8).

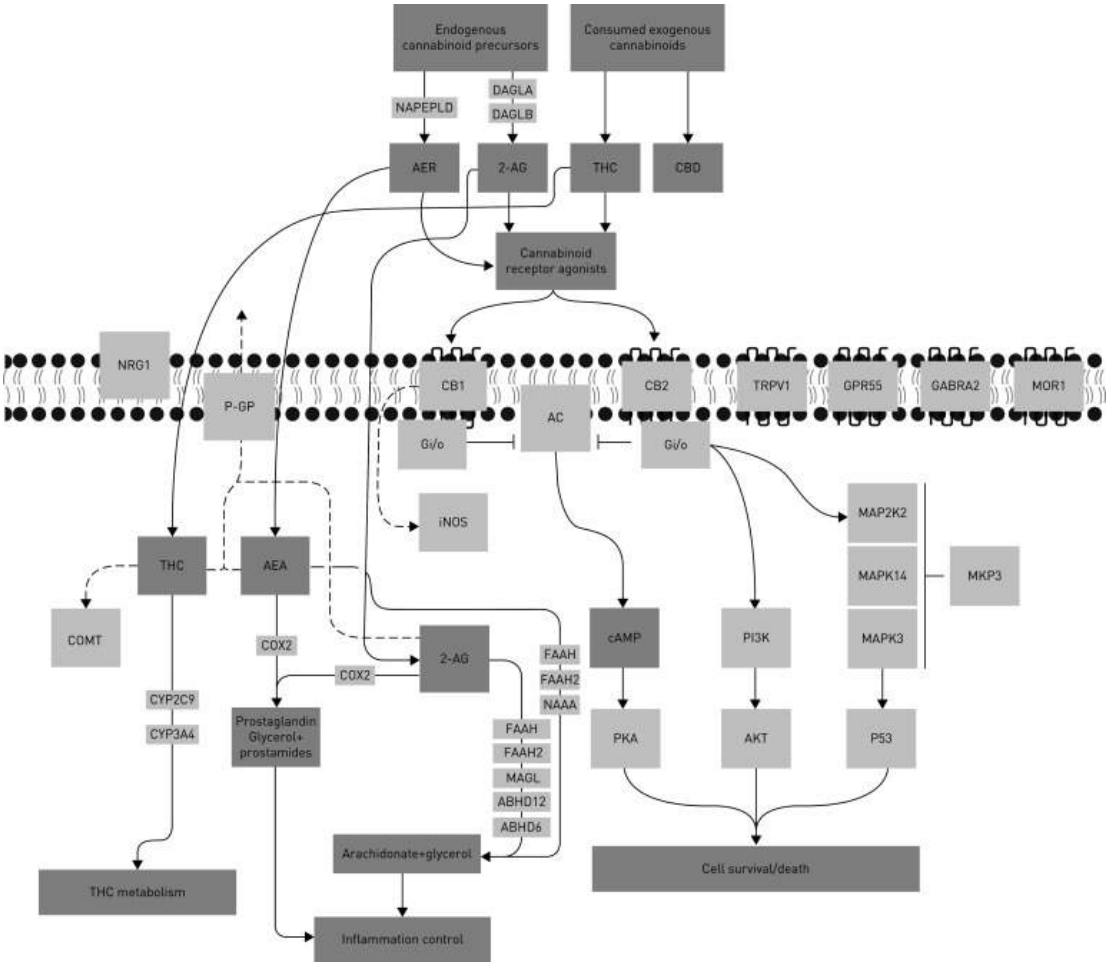


Figure 3.6: Solid arrows indicate known relationship between candidates and ligands. Dotted arrows indicate proposed relationships. Blunted lines indicate inhibition. Candidate functions are annotated in Table 8. THC: tetrahydrocannabinol.

To begin the characterization of the endocannabinoid system in human airway epithelial cells, *in situ* localization of CB1, CB2, and TRPV1 protein in human lung tissue was performed (Figure 3.7). It is demonstrated that all three receptors are present at protein level in human airway epithelial cells *in situ* relative to negative control (Figure 3.7 A, C, E, G). To validate the staining patterns observed and antibody specificity, immunoblots with primary human airway epithelial cells, whole lung tissue, and PBMCs were performed. A single band for CB1 was observed at approximately 45 kDa in human airway epithelial cells, but not whole human lung, or PBMCs (Figure 3.7 B). A dominant band for CB2 was observed at approximately 40 kDa and accompanied by a reported 52-55 kDa doublet[436] in human airway epithelial cells, with a similar pattern observed for PBMCs (Figure 3.7 D). In contrast, in whole human lung the dominant band was observed at 55 kDa with only a faint 40 kDa band. The band patterns observed for CB2 are consistent with glycosylation of the N-terminus and processing of the full-length peptide.[436] A dominant band for TRPV1 was observed at approximately 100 kDa in human airway epithelial cells and accompanied by two lower molecular weight bands at approximately 70 kDa and 37 kDa (Figure 3.7 F). No TRPV1 bands were observed in whole human lung, while a dominant single band at 100 kDa was observed in peripheral blood mononuclear cells. Total protein loading staining from a representative blot demonstrates equal loading within replicates of the same sample type and distinct protein compositions between sample types (Figure 3.7 H).

To extend the *in situ* CB1 and CB2 protein staining and to interrogate the potential of non-specific staining generated with the CB2 antibody, collaborators next performed *in situ* hybridisation for *CNR1* and *CNR2* gene transcripts using RNAscope® technology (Figure 3.8). *In situ* hybridisation demonstrates transcripts for both *CNR1* (Figure 3.8 A) and *CNR2* (Figure 3.8 B) in the airway epithelium of all human lung samples examined (n = 10). Serial sections from each lung sample reveal that *CNR1* and *CNR2* are likely to co-express in airway epithelium.

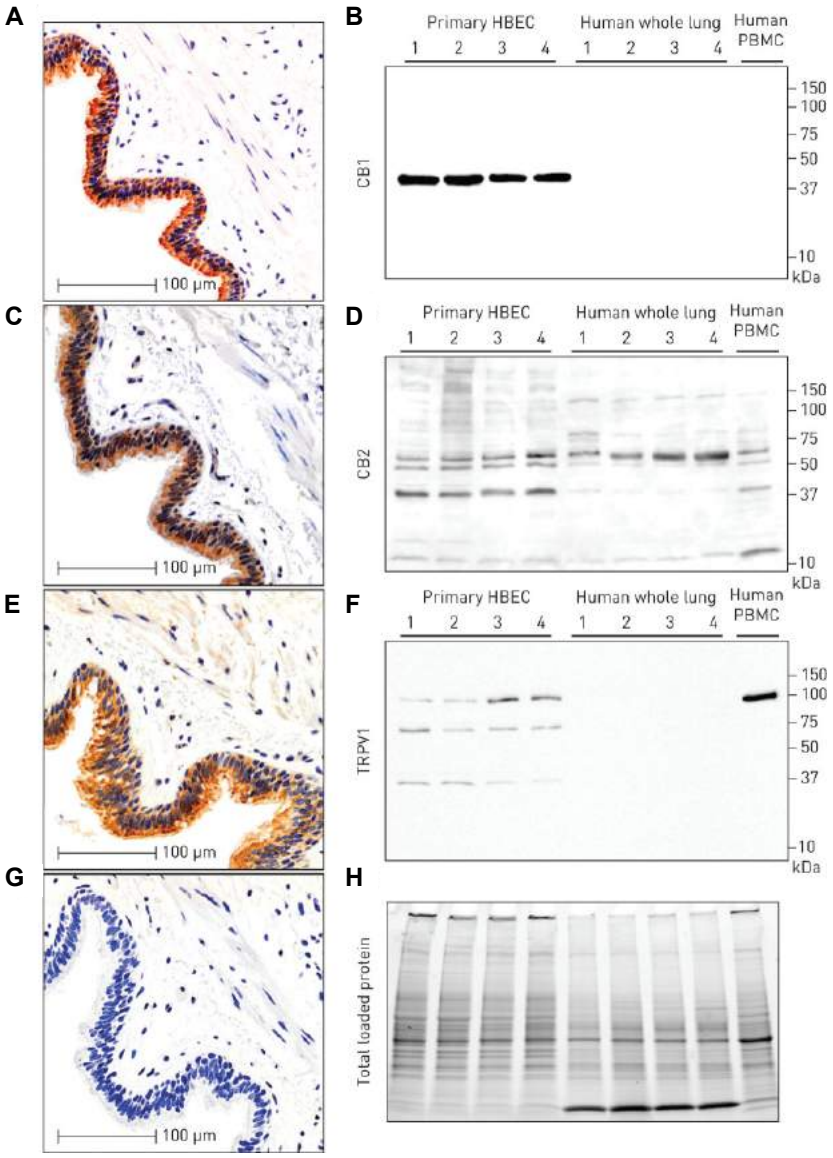


Figure 3.7: Serial sections from a single patient donor that is representative of $n = 10$, for immunohistochemistry of (A) CB1, (C) CB2 and (E) TRPV1 with (G) negative control. Immunoblots on primary human airway epithelial cells cultured *in vitro*: (B) CB1, (D) CB2, and (F) TRPV1 with (H) total protein loading control ($n = 4$ airway epithelial cells (HBEC), $n = 4$ whole-lung samples, $n = 1$ peripheral blood mononuclear cells (PBMCs)). Molecular weights (in kilodaltons) are denoted on y-axis of immunoblots.

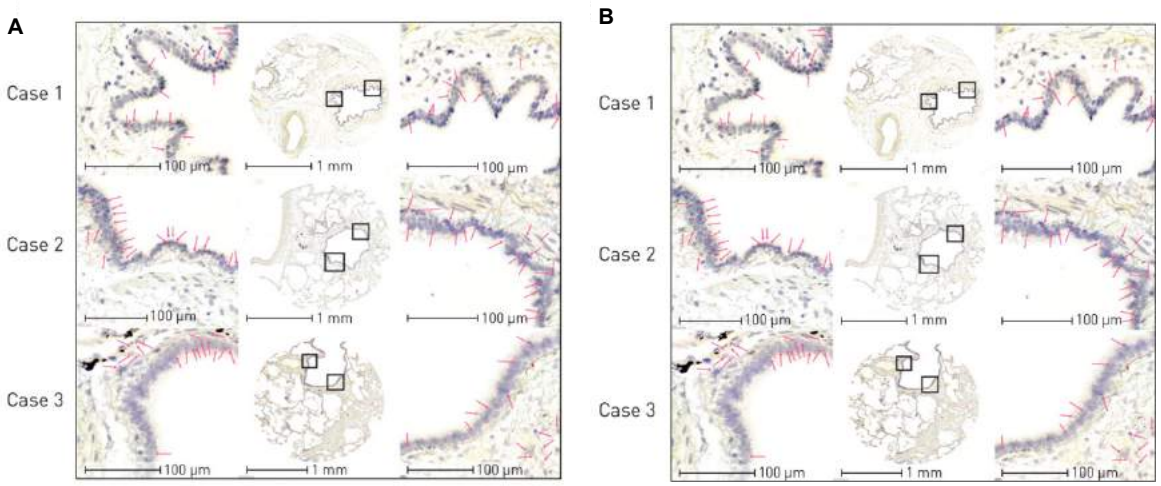


Figure 3.8: *In situ* hybridisation of (A) *CNR1* and (B) *CNR2* in three patient donors representative of $n = 10$. Serial sections of each of the three cases were stained for *CNR1* and *CNR2*. Low magnification images are in the centre of (A) and (B), with high-powered magnification regions of interest on either side highlighted by black boxes. Pink arrows correspond to positive puncta representative of mRNA transcript.

To further corroborate the *in situ* CB1 and CB2 protein and gene transcript staining, promoter activity data for *CNR1* and *CNR2* were extracted and analysed from the FANTOM5 data set, which includes 1886 primary cells, cell lines, and tissue sample types. All samples that included "lung", "nasal", "trachea", "bronchial", "airway", or "alveolar" were selected to identify lung-specific sample types ($n = 28$). *CDH1* promoter activity was used as a positive control. Consistent with the observed gene expression analysis in airway epithelial cells, normalized transcripts per million (TPM) values for each CAGE peak demonstrate that *CNR1* and *CNR2* promoter activity was present but modest across airway epithelial cells and lung tissue samples (Figure 2). Both microarray gene expression analysis and promoter activity were consistent with results of candidate gene and protein expression in airway epithelial cells via *in situ* hybridisation and immunohistochemistry.

Collectively, the *in situ* and *in vitro* data confirm expression of CB1, CB2, and TRPV1 protein in human airway epithelial cells.

Sex differences in endocannabinoid system gene expression in human airway epithelial cells from healthy subjects

Sex differences in CB1 and CB2 expression levels have been reported,[304] which could impact downstream responses to cannabinoid exposures. Furthermore, sex and gender differences in cannabis consumption practices have been reported.[68–70] Collectively, these two factors could interact and contribute to differential responses to cannabinoids in distinct populations.

To examine sex differences in the endocannabinoid system, I analysed the 32-gene endocannabinoid signature in a curated data set of airway epithelial cells from 616 unique healthy subjects, where the identifier of sex was available for 103 females and 227 males (Table 9). The 32-gene endocannabinoid signature was first analysed in all 616 subjects to show overall trends for each gene (Figure 3.9 A). Subsequently, a PCA plot was performed for all samples with sex as an identifier (Figure 3.9 B). The PCA plot reveals clustering of samples from males within a larger space occupied by the female samples. Statistical analysis at the individual gene level revealed a difference in 7 of 32 genes (Figure 3.9 C). Five genes (*ABHD6*, *MAPK14*, *NAAA*, *NRG1*, and *PIK3CA*) were down-regulated in males relative to females, while two genes (*CYP2C9* and *GPR55*) were up-regulated in males relative to females. The gene expression patterns were overlaid on the endocannabinoid signalling pathway for qualitative visualization of global changes in the 32-gene endocannabinoid signature in males relative to females (Figure 3.9 D).

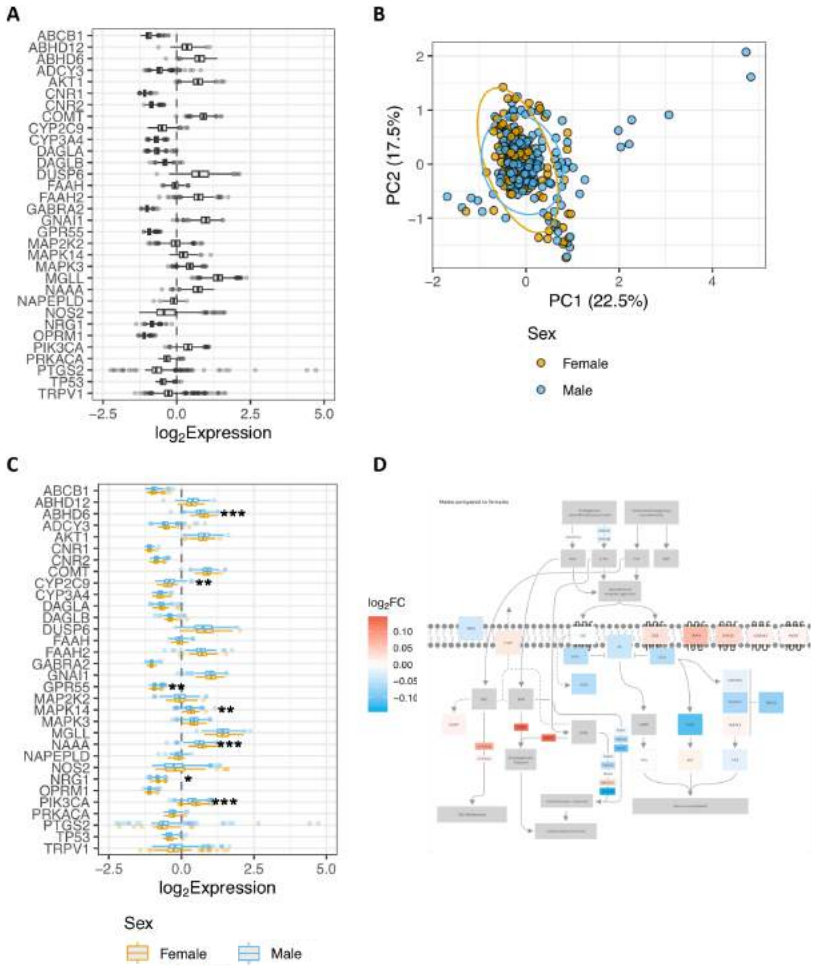


Figure 3.9: (A) Gene expression data for 616 healthy subjects with no history of smoking or chronic respiratory disease. (B) Principal-component (PC) analysis plot of healthy females ($n = 103$) and males ($n = 227$) generated by expression patterns of the 32-gene endocannabinoid signature. The first (22.5%) and second (17.5%) PCs were used. Ellipses were added to represent 95% confidence intervals per sex. (C) Healthy samples with metadata defining sex were further divided into male and female groups and plotted separately as blue and orange-outlined box plots, respectively. For both (A) and (C), $\log_2$-transformed expression values were plotted as box plots. The dashed line at zero represents the global baseline of expression for the entire set of genes. (D) Visual representation of the differences between healthy females and males in the 32-gene endocannabinoid signature. Colour coding is reflective of $\log_2$ fold change (FC) of males relative to females. THC: tetrahydrocannabinol. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Collectively, the data in this study confirm that the endocannabinoid system is expressed at gene level in human airway epithelial cells, suggesting that signaling downstream of receptors is intact. Importantly, sex differences are observed in the 32-gene endocannabinoid signature expression in healthy subjects, with 5/7 differentially abundant genes exhibiting higher gene expression in females relative to males.

The endocannabinoid system is dysregulated in human airway epithelial cells from individuals with asthma and COPD

In addition to sex, disease status may also impact the expression of the endocannabinoid system in airway epithelial cells, as specific phenotypes are observed in cells isolated from asthmatics and individuals with COPD.[208,222,364] I therefore tested the hypothesis that the 32-gene endocannabinoid signature was dysregulated in asthma and COPD.

To test this hypothesis, I curated all 1090 samples that included 616 healthy subjects, 136 subjects with asthma, and 338 subjects with COPD. A PCA plot reveals clustering of samples from healthy, asthmatic, and COPD subjects, with samples from people with asthma separating from both healthy subjects and those with COPD (Figure 3.10 A). Statistical analysis at the individual gene level revealed changes in 21 of 32 genes in people with asthma and 26 of 32 genes in those with COPD (Figure 3.10 B). In people with asthma, 11 of 21 dysregulated genes were up-regulated (*ABCB1*, *ABHD6*, *CYP2C9*, *GABRA2*, *GNAI1*, *GPR55*, *NOS2*, *NRG1*, *PTGS2*, *TP53*, and *TRPV1*), while 10 of 21 were down-regulated (*ABHD12*, *CNR2*, *COMT*, *CYP3A4*, *DAGLA*, *DAGLB*, *MAP2K2*, *MAPK3*, *MGLL*, and *NAAA*). In COPD subjects, 11 of 26 dysregulated genes were up-regulated (*ABHD6*, *CNR1*, *CYP2C9*, *CYP3A4*, *FAAH*, *GABRA2*, *GPR55*, *NAAA*, *NAPEPLD*, *OPMRI*, and *TRPV1*), while 15 of 26 were down-regulated (*ABCB1*, *ABHD12*, *ABCY3*, *AKT1*, *DAGLB*, *DUSP6*, *FAAH2*, *GNAI1*, *MAPK14*, *MAPK3*, *MGLL*, *NOS2*, *PIK3CA*, *PRKACA*, and *TP53*). The most dysregulated gene was *TRPV1*, with the largest up-regulation observed in the samples from people with asthma relative to healthy controls. The differential gene expression patterns were overlaid on the endocannabinoid signalling pathway for qualitative visualization in asthma (Figure 3.10 C) and COPD (Figure 3.10 D).

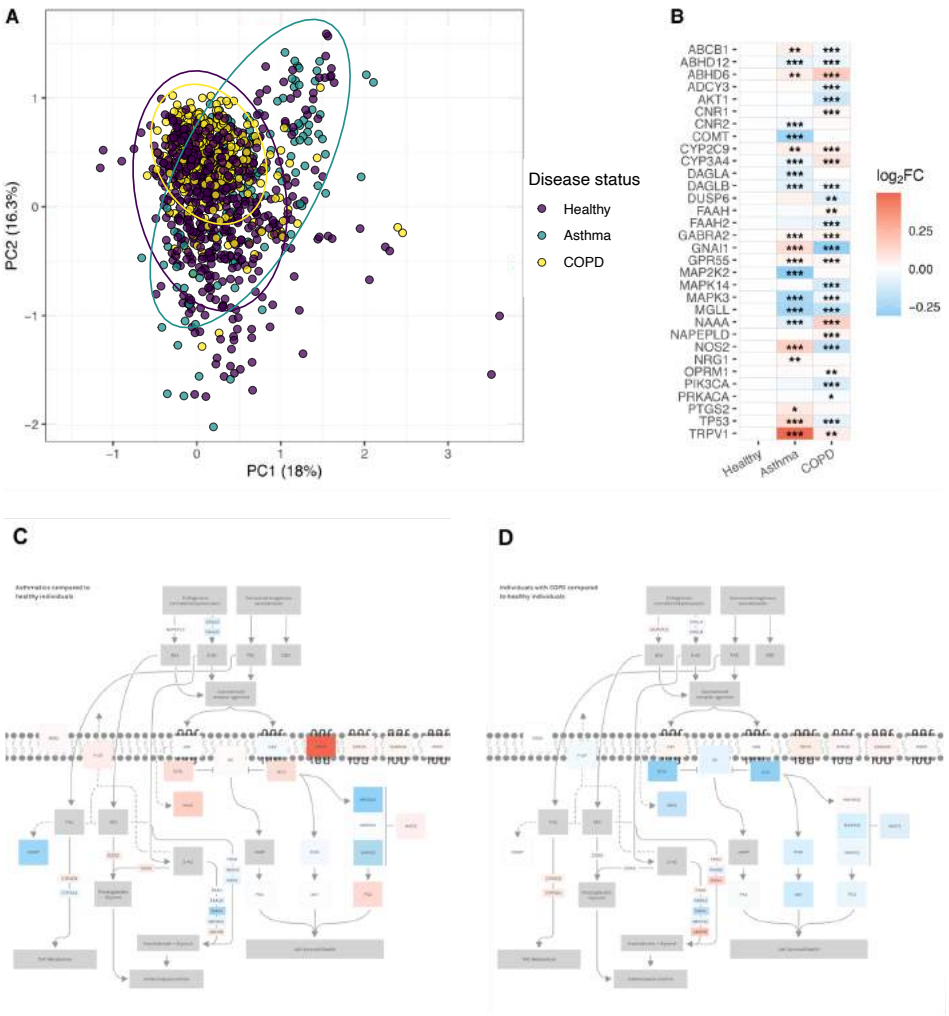


Figure 3.10: (A) Principal-component (PC) analysis plot of healthy subjects ($n = 616$), people with asthma ($n = 136$) and individuals with COPD ($n = 338$) generated by expression patterns of the 32-gene endocannabinoid signature. The first (18%) and second (16.3%) PCs were used. Ellipses were added to represent 95% confidence intervals per sex. (B) Gene expression data of the 32 genes were compared between healthy, asthmatic and COPD samples. The $\log_2$-transformed mean expression values were compared to that of the healthy samples and shown as $\log_2$ fold change (FC). Visual representation of the differences in the 32-gene endocannabinoid signature between (C) healthy subjects and people with asthma and (D) healthy subjects and individuals with COPD. Colour coding is reflective of $\log_2$FC relative to healthy subjects. THC: tetrahydrocannabinol. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Collectively, the data in this study demonstrate that underlying chronic respiratory disease status is associated with a dysregulation of the endocannabinoid system at the gene level in human airway epithelial cells.

Impact of sex on endocannabinoid system gene expression in human airway epithelial cells from asthmatics and subjects with COPD

Sex differences in incidence, age of onset, and pathology are observed in both asthma and COPD.[306,358] Cannabinoid exposures have been explored in the context of both asthma and COPD management for immunomodulatory and bronchodilation purposes.[5,377,379] To date, the potential interaction of sex status and endocannabinoid system expression in chronic respiratory disease has not been addressed.

Taking the same approach as for healthy subjects, I analyzed a curated data set of airway epithelial cells from 136 unique asthmatic subjects, where the identifier of sex was available for 34 females and 28 males. For COPD, I analyzed a curated data set from 338 unique COPD subjects, where the identifier of sex was available for 48 females and 93 males (see Table 9 for study group compositions).

In both asthma and COPD samples, PCA plots revealed no separation between sexes with clustering of samples overlapping between disease groups (Figure 3.11 A, C). At the individual gene level, no sex dependent differences were observed for any gene in the 32-gene endocannabinoid signature in either disease groups (Figure 3.11 B, D).

Collectively, the data do not support sex differences in the endocannabinoid system in human airway epithelial cells from asthmatics or subjects with COPD.

TRPV1 is up-regulated in airway epithelial cells from asthmatics

As my bioinformatic interrogation of the 32-gene endocannabinoid signature was restricted to genes, confirmatory protein expression analysis was performed. The candidate chosen for validation was TRPV1, a confirmed receptor for cannabinoids that was the most differentially expressed candidate between the comparisons examining sex or disease status.

Using primary human airway epithelial cells from healthy donors or those with physician diagnosed asthma, protein was isolated from cells grown under submerged monolayer culture conditions. Immunoblot analysis confirms TRPV1 protein expression in human airway epithelial cells and revealed a qualitative increase in cells from asthmatics (Figure 3.12). TRPV1 staining normalized to total protein loading confirms a quantitative increase. In

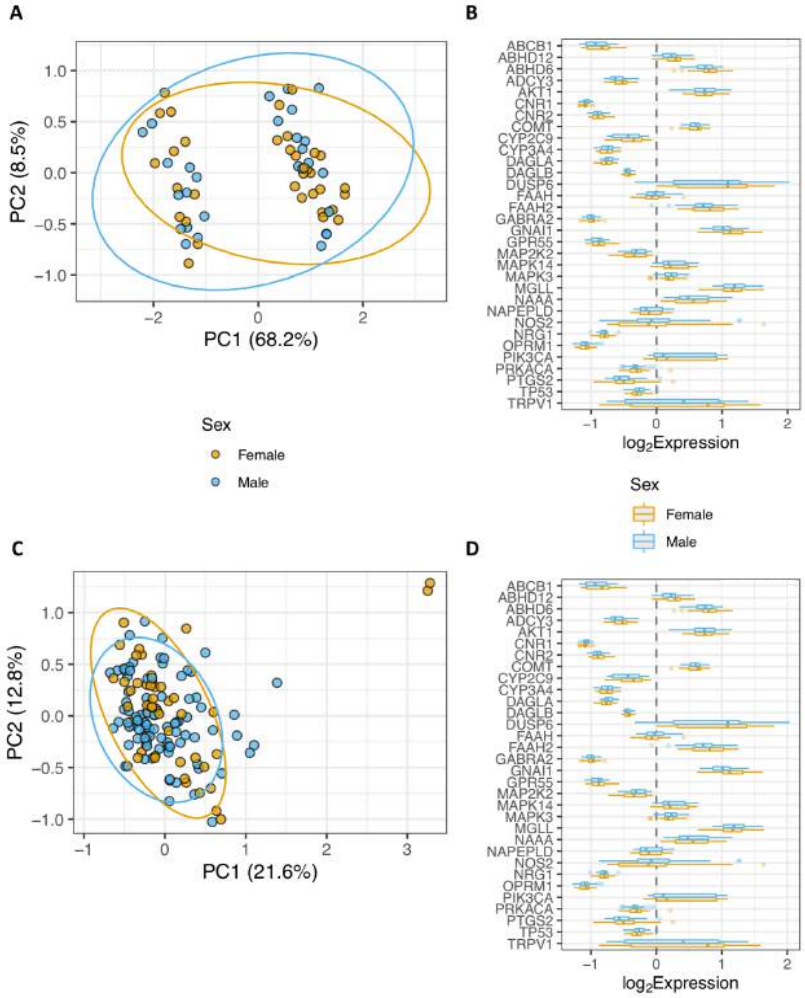


Figure 3.11: (A) Principal-component analysis (PCA) plot of asthmatic females ($n = 34$) and males ($n = 28$) generated by expression patterns of the 32-gene endocannabinoid signature. The first (68.2%) and second (8.5%) principal components (PCs) were used. (B) Asthmatic samples divided into female and male, and plotted separately. (C) PCA plot of females ($n = 48$) and males ($n = 93$) with COPD generated by expression patterns of the 32-gene endocannabinoid signature. PC1 (21.6%) and PC2 (12.8%) used. (D) COPD samples divided into female and male, and plotted. For both (A) and (C), ellipses were added to represent 95% confidence intervals per sex. For both (B) and (D), $\log_2$-transformed expression values were plotted as box plots of $\log_2$ fold change (FC). The dashed line at zero represents the global baseline of expression for the entire set of genes.

closing, the performed protein analysis is consistent with the bioinformatic analysis that revealed elevations in TRPV1 in human airway epithelial cells from people with asthma.

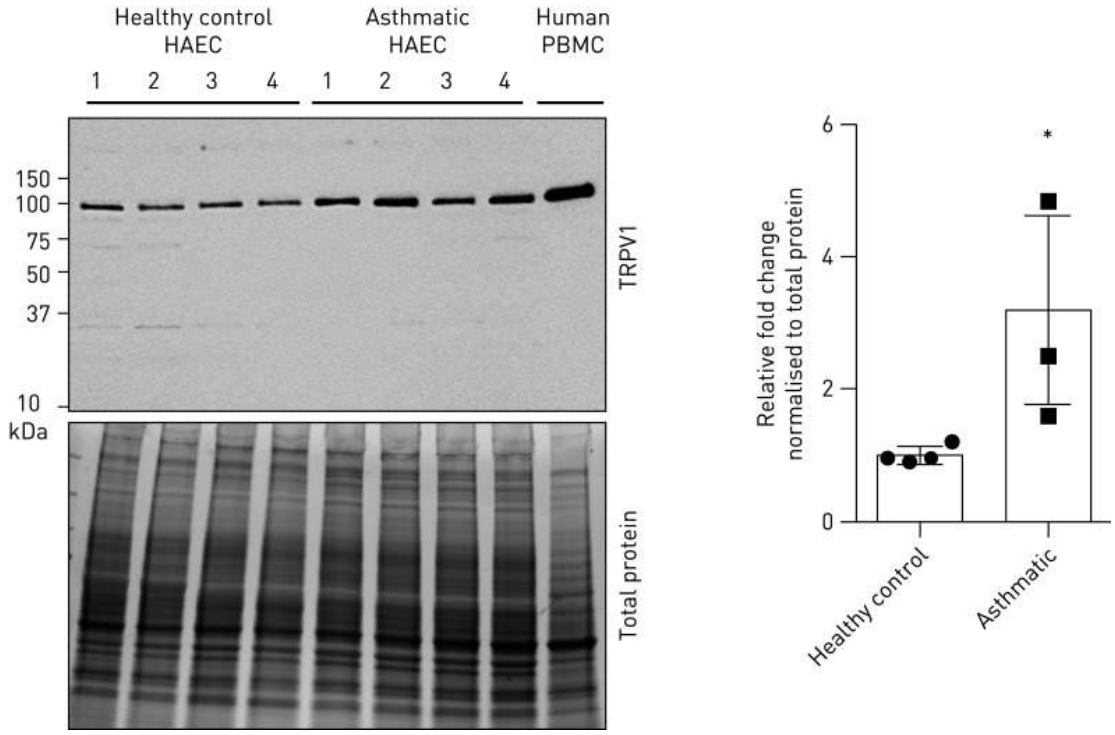


Figure 3.12: Immunoblot of HAECs from healthy subjects ($n = 4$), people with asthma ($n = 4$) and peripheral blood mononuclear cell (PBMCs) (control) were analysed for TRPV1 and quantified as fold change over healthy subjects, normalized to total protein loading. *: $p < 0.05$.

3.3.4 Discussion

The dominant route of delivery of cannabis is via inhalation of combustion smoke, resulting in exposure to phytocannabinoids and activation of the endocannabinoid system.[68–70] The airway epithelium represents the first line of defence in the human lung against inhaled insults including cannabis smoke. To better understand how the epithelium is able to respond to cannabis smoke exposure in the context of the endocannabinoid system, I performed a characterization study using bioinformatic and complementary protein analysis approaches. This study demonstrates that three cannabinoid receptors, CB1, CB2, and

TRPV1, are expressed at the protein level in human airway epithelium *in situ* and *in vitro*. The demonstration that these receptors were present warranted an exploration into the endocannabinoid system downstream of the receptors. Using 1090 unique patient samples of airway epithelial cells curated from publicly available data sets, I demonstrate in healthy subjects that the gene expression levels of the endocannabinoid system are elevated in females relative to males. Furthermore, I demonstrate that the sex-dependent effect is lost when examining samples from asthmatics or individuals with COPD, and that disease status appears to be a stronger driver of endocannabinoid system gene expression. Lastly, my bioinformatic approach was validated by demonstrating that TRPV1, a top candidate up-regulated at the gene level in these studies, is also up-regulated at the protein level in asthmatics. Collectively, these results confirm the expression of the endocannabinoid system in human airway epithelial cells and suggest that both sex and disease status may impact cannabinoid responses.

The legalization of cannabis in multiple jurisdictions on a global scale reduces barriers for individuals to consume cannabis for either medicinal or recreational purposes. The dominant route of cannabis delivery is through inhalation of smoke from plant combustion.[68–70] Inhaled cannabis smoke travels through the upper and lower airways, with airway epithelial cells being a major site of first contact. This study has demonstrated with an *in vitro* model of airway epithelial cell culture that cannabis smoke induces a concentration-dependent reduction in airway epithelial cell viability, barrier function, while promoting pro-inflammatory cytokine secretion.[10,178] Complementary profiling of human epithelial cells isolated via bronchial brushings has demonstrated cannabis consumption-dependent elevations in *TLR5*, *TLR6*, and *TLR9* gene expression.[31] A limitation of these studies is the lack of mechanistic interrogation into the role that the endocannabinoid system contributed to the observed functional consequences of cannabis smoke. To begin to implicate the endocannabinoid system in epithelial cell functions, an *in vitro* approach has used direct cannabinoid administration to airway epithelial cells independent of combustion, showing functional consequences with altered barrier function mediated by mechanisms dependent and independent of cannabinoid receptors.[342,343] The demonstration that multiple components of the endocannabinoid system are expressed at the gene level in airway epithelial cells with CB1, CB2, and TRPV1 receptors confirmed at the protein level are consistent with this system playing a contributing role in mediating cannabis smoke-induced effects. Collectively, the data in this study and existing literature demonstrate that the dominant form of cannabis consumption, cannabis smoke inhalation, is able to induce functional consequences in airway epithelial cells in a process that is mediated, at least in part, by the endocannabinoid system.

Cannabis use patterns have been reported to differ among males and females, with

females consuming less overall quantity and frequency.[68–70] The endocannabinoid system is in turn regulated by sex hormones with diverse interactions at levels of receptors, enzymes, and signaling molecules.[29,285] In the context of lung health and disease, sex dependent lung physiology is observed as well as asthma and COPD disease incidence and progression.[306,358] In light of the potential for interactions between sex, user practices, and the endocannabinoid system, I performed a bioinformatic analysis that stratified for male and female sex status. In healthy individuals, I observed that 7/32 of the endocannabinoid gene signature candidates were differentially regulated between males and females, with 5/7 genes being up-regulated in females. The observations of sex-dependent effects on the expression of the 32-gene endocannabinoid signature are consistent with sex differences in the expression of CB1 and CB2 protein expression in heart tissue, where CB1 receptors are more highly expressed in females and CB2 receptors more highly expressed in males.[304] Interestingly, age confounded these results in heart tissue with a switch in expression patterns observed over the ages of 50. Despite the observed sex-differences in healthy subjects, these observations were not observed in samples from asthmatics or individuals with COPD, where instead, disease status was associated with differential expression patterns. Although both asthma and COPD are influenced by sex, with sex hormones impacting airway epithelial cell biology,[369] these results suggest that asthma or COPD status masks any sex-dependent effect. Collectively, these findings suggest that sex and asthma or COPD status independently impact expression of the endocannabinoid system.

The potential disease specific responses to cannabis exposure as a result of dysregulated endocannabinoid signaling is relevant as cannabinoids have and are being pursued for bronchodilatory and immunomodulatory properties.[5,212,377,379] Furthermore, in populations of asthmatics and COPD subjects, cannabis consumption is not avoided and shown to interact with disease progression. The reported benefits of cannabis exposure on lung function in asthmatics may be selective for this population based on endocannabinoid signaling. Indeed, if TRPV1 is a dominant receptor for responses downstream of inhaled cannabis, the signaling mediated in airway epithelial cells from asthmatics may be augmented relative to healthy controls. The observation of elevated TRPV1 in airway epithelial cells from asthmatics is consistent with a previous report demonstrating correlation with asthma severity.[258] A limitation of this study is that other cell-types were not explored for TRPV1, which is also expressed in airway smooth muscle cells and can modulate smooth muscle contraction.[431] In contrast, the lack of observed benefit on lung function and association with advanced pathology in COPD subjects may be a result of a distinct expression profile of the endocannabinoid system.[5,117,371] These results and those in the literature suggest that a universal response to inhaled cannabis by healthy subjects and those with asthma or COPD should not be assumed and cautions translation of safety and efficacy studies

performed in healthy individuals to those with underlying asthma or COPD.

This study is heavily focused on using deposited data sets generated by microarray gene expression technology. Microarrays are low cost, cover large numbers of transcripts, and benefit from a standardized format for analysis and public deposition of data. As a result of these benefits, gene expression data for large, independent cohorts are available for curation and data mining purposes if appropriate measures are taken to normalize and integrate data sets from diverse user groups. Despite benefits of curated microarray data sets, there exist limitations of such an approach. Specifically, microarray technology is susceptible to poor signal-to-noise ratio, making transcripts expressed at low levels difficult to detect, suggesting that the signal may be under-estimated or high variance may be introduced for these transcripts. To address this limitation, the SCAN method used functions to reduce technical and across-sample variation and to increase the signal-to-noise ratio while maintaining the ability to detect differentially expressed genes.[301] As in all analyses, the sample size will dictate the statistical power, which can only be as large as the available data sets. In this case, the sample sizes for females and males in the asthma and COPD data sets were smaller than the healthy cohorts, which may have limited my ability to detect any sex-specific effects on endocannabinoid gene expression. Finally, in order to compare data across multiple experiments the microarray chip technology used needs to be considered with additional normalization methods required for data generated on different platforms. Different microarray technologies will have unique probe compositions and quality control protocols which can lead to systematic biases between experiments, though these batch-specific effects can be addressed using batch correction techniques and normalization.[249] Despite these highlighted limitations, the curation of multiple microarray gene expression data sets from diverse cohorts of study subjects can be aligned with normalization methods to minimize batch and cross-platform effects and maximizing sample sizes to detect differential gene expression patterns.

In summary, this study demonstrates that the endocannabinoid system is expressed in human airway epithelial cells at the gene level and CB1, CB2, and TRPV1 are expressed at the protein level. I demonstrate that gene expression patterns for a 32-gene endocannabinoid signature are differentially expressed between females and males in healthy individuals and between healthy individuals and those with asthma or COPD. It can be confirmed that my bioinformatic approach for analysis of gene expression has the potential to reflect corresponding protein expression level changes as demonstrated by elevated TRPV1 protein expression in human airway epithelial cells from asthmatics relative to healthy controls. This study lays a foundation with primary human lung samples from well defined patient populations to justify exploring the functional consequences of endocannabinoid system signaling in human airway epithelial cells in both health and disease. The complete functions

of the endocannabinoid system in airway epithelial cells remain to be defined.

3.4 Summary

Cannabis and tobacco smoke exposures produce strikingly similar transcriptome-wide outcomes in the human airway epithelium, including increased oxidative stress, increased pro-inflammatory responses, and a shared response to long-acting β-agonist/glucocorticoid therapy.[10,178] However, the long-term effects of chronic cannabis use are still not well-established. This is due, in part, to a paucity of large-scale sequencing analyses and lack of a standardized product. Differences in the expression of endocannabinoid receptors between healthy males and females and between healthy individuals and those who are diagnosed with asthma or COPD suggest variability between these groups with respect to long-term outcomes, further necessitating large-scale, sequencing-based, longitudinal studies into cannabis use and it's effects on the human airway.

A common suggestion is for cannabis users to switch from combustion to vaporization as their primary mode of delivery however, the effects of vaporizing cannabis are also poorly understood despite consumption via vaporization increasing in popularity from 34 to 40% between 2017 and 2018.[68,69] Any longitudinal studies into cannabis use should include various delivery routes with enough participants in each sample group to ensure sufficient statistical power. This previous data chapter has highlighted the vast capabilities of current computational methods in answering biologically significant questions however, these methods are dependent on the availability of large, well-annotated sample groups and deep-sequencing in order to do so.

It is evident that cannabis use poses a major quandary to global public health and will continue to be a pertinent area of research in the years to come. It is imperative that appropriate funding be allocated to such research areas in order to fully benefit from what current bioinformatics has to offer. Bioinformatics allows researchers to be on the cutting edge, responding rapidly to pressing public health concerns, be that cannabis consumption or something far more novel.

Chapter 4

SARS-CoV-2 infection

> We won't choose whether the next novel disease pandemic emerges, but we will choose how ready we are and how quickly and equitably we respond.
>
> — John Green

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in December 2019 to cause a pandemic of coronavirus disease (COVID-19).[440] SARS-CoV-2 is spreading at a much more rapid rate than SARS-CoV.[76,269,440] In 2003, the SARS outbreak caused by the SARS coronavirus (SARS-CoV) resulted in 8096 probable cases with 774 confirmed deaths.[41,118] In patients with SARS, deaths were attributed to acute respiratory distress associated with diffuse bilateral pneumonia and alveolar damage.[296] Similar clinical reports of diffuse bilateral pneumonia and alveolar damage have been reported for SARS-CoV-2 infection.[142,421,445] SARS-CoV-2 causes a respiratory infection, along with acute respiratory distress syndrome in severe cases. Severe cases of SARS-CoV-2 have been associated with infections of the lower respiratory tract, with detection of the virus throughout this tissue as well as the upper respiratory tract.[142,421,445] The biological mechanisms that may govern differences in the number of SARS and COVID-19 cases remain undefined. It is possible that SARS-CoV-2 possesses distinct molecular mechanisms that affect the virulence through viral proteins, greater susceptibility of host cells to infection, permissivity of host cells to virus replication, or some combination of these and other potentially unknown factors.[232,245,419,427] Understanding SARS-CoV and SARS-CoV-2 virus similarities and differences at the molecular level in the host may provide insights into transmission, pathogenesis and interventions.

As well as host-viral entry mechanism, characterization of the host-viral immune response

is critical for guiding public health measures and intervention. Type I interferons (IFNs) are one of our first lines of defence against a virus. Protein over-expression studies have suggested the ability of SARS-CoV-2 proteins to block IFN responses. Emerging data also suggest that timing and extent of IFN production is associated with manifestation of COVID-19 severity. In spite of progress in understanding how SARS-CoV-2 activates antiviral responses, mechanistic studies into wild-type SARS-CoV-2-mediated induction and inhibition of human type I IFN responses are lacking.

This chapter aims to characterize the mechanisms facilitating SARS-CoV-2 viral entry into human airway cells and to assess the various host responses observable once viral entry is achieved. The data suggest that dynamic regulation of ACE2 expression in human lung and/or the existence of alternate receptors for SARS-CoV-2 facilitate initial host cell infection in lung tissue. Once the virus has entered the host cell, it is demonstrated that SARS-CoV-2 infection induces a mild type I IFN response *in vitro* and that SARS-CoV-2 is not adept in blocking these responses, providing support for ongoing IFN clinical trials. This research expands our current knowledge base and may inform public health efforts in the handling of both current and emergent zoonotic pathogen pandemics.

4.1 Gene expression and *in situ* protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue

Material in this section has been published as part of Aguiar *et al.* (2020).[12] The published manuscript is available here:

> J. A. Aguiar, B. J.-M. Tremblay, M. J. Mansfield, O. Woody, B. Lobb, A. Banerjee, A. Chandiramohan, N. Tiessen, Q. T. Cao, A. Dvorkin-Gheva, S. Revill, M. S. Miller, C. Carlsten, L. Organ, C. Joseph, A. John, P. Hanson, R. C. Ausin, B. M. McManus, G. Jenkins, K. Mossman, K. Ask, A. C. Doxey, and J. A. Hirota. Gene expression and *in situ* protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue. *European Respiratory Journal*, 56(3):2001123. 2020.[12] https://doi.org/10.1183/13993003.01123-2020

4.1.1 Introduction

The seminal report identifying the receptor for SARS-CoV used a HEK293 cell over-expression system to identify angiotensin-converting enzyme 2 (ACE2) as a receptor by co-immunoprecipitation with the SARS-CoV spike domain 1.[238] Subsequently, the spike (S) protein of SARS-CoV was identified as the viral interacting partner of ACE2. Host protease activity by transmembrane serine protease 2 (TMPRSS2) was then shown to facilitate ACE2 ectodomain cleavage and fusion of the SARS-CoV membrane with the host cell membrane.[85,131,349]

ACE2 and TMPRSS2 were identified as cellular entry determinants for SARS-CoV using mechanistic studies. The original report of *in situ* human lung ACE2 expression described positive immunohistochemical staining for alveoli and airway epithelial cells, and immunocytochemical staining in A549 type II alveolar epithelial cells.[149] ACE2 protein expression is also present in the human lung adenocarcinoma cell line Calu-3.[385] Similar to ACE2, the original report describing the expression of TMPRSS2 in human respiratory mucosa described expression in airway epithelium and type II alveolar epithelial cells.[46] The specificity of the ACE2 and TMPRSS2 antibodies used for analysis of expression patterns in human lung tissues remains to be addressed.

In addition to TMPRSS2, ADAM17 (ADAM metallopeptidase domain 17, a member of the ADAM (a disintegrin and metalloprotease domain) family), has also been demonstrated to cleave ACE2 ectodomain, but this was not required for SARS-CoV infection.[148,158,187] Mechanisms of SARS-CoV entry distinct from ACE2 have also been reported and include activation by endosomal cathepsin L and cell surface expression of CD147 (also known as basigin (BSG)) or GRP78 (78 kDa glucose-regulated protein; also known as heat shock protein family A (Hsp70) member 5 (HSPA5)).[86,88,350] Each of these receptors was mechanistically interrogated and results suggested that SARS-CoV could initiate host cell entry and infection using multiple mechanisms.

Recent *in vitro* reports have demonstrated that similar host proteins are involved in facilitating cell entry by SARS-CoV-2, such as ACE2 and TMPRSS2.[169,440] Biophysical and structural evidence strongly supports an interaction of ACE2 with SARS-CoV-2 spike protein, similar to SARS-CoV spike protein.[419,426] Molecular docking studies have also suggested that SARS-CoV-2 spike protein can interact with cell surface GRP78.[185] Indirect evidence for a role of CD147 in SARS-CoV-2 binding has been demonstrated *in vitro* with the use of an anti-CD147 intervention that prevented virus replication.[407] Furthermore, a clinical study with an anti-CD147 intervention reduced symptoms and duration of hospital admission for COVID-19 patients.[49] In summary, although there is evidence that SARS-CoV-2 and SARS-CoV both utilise ACE2 as a receptor to facilitate virus entry, it is possible

that differences in host entry mechanisms play a role in the large epidemiological differences between the two viruses, which may include additional unidentified receptors.

To address the uncertainties related to SARS-CoV-2 receptors in human lung, gene expression and *in situ* protein profiling of candidate receptors in human airway epithelial cells and lung tissue were performed. My computational analysis used publicly available microarray gene expression data sets from airway epithelial cells of 515 unique subjects, single cell sequencing data from 10 subjects, and the FANTOM5 data set for promoter activities of 74 lung-related cell and tissue types. My proteomic analysis used data from the Human Proteome Map[211] and a data set from primary human airway epithelial cells grown under air–liquid interface culture conditions.[119] For the *in situ* protein profiling, immunohistochemical analysis of 98 human lung tissue samples was performed. To determine antibody specificity, collaborators performed immunoblots on protein isolated from Calu-3 cells, primary human airway epithelial cells, primary type II alveolar epithelial cells, the human bronchial epithelium cell (HBEC)-6KT cell line, the A549 type II alveolar epithelial cell line, and HEK cells. Collectively, the data contrast with previous reports, demonstrating rare ACE2 protein expression in the airway epithelium and alveoli of human lung. The protein expression data are consistent with low ACE2 promoter activity in a variety of lung epithelial cell samples and low ACE2 gene expression in both microarray and single cell RNA sequencing (scRNA-seq) data sets. This study presents confirmatory evidence for the presence of TMPRSS2, CD147 and GRP78 protein *in vitro* in airway epithelial cells and confirm broad *in situ* protein expression of CD147 and GRP78 in the respiratory mucosa. The data suggest the presence of a mechanism dynamically regulating ACE2 expression in human lung, perhaps in periods of SARS-CoV-2 infection, and/or that alternate receptors for SARS-CoV-2 exist to facilitate initial host cell infection in lung tissue.

4.1.2 Methods

Human ethics

Procurement of primary human airway epithelial cells used for immunoblots and lung tissue for immunohistochemistry was approved by the Hamilton (ON, Canada) integrated Research Ethics Board (HiREB 5099T, 5305T, 11-3559 and 13-523-C). The University of British Columbia (Vancouver, BC, Canada) Research Ethics Office approved heart tissue archives and primary human airway epithelial cell collection.

Upper and lower airway gene expression analysis

Public microarray experiments using Affymetrix chips (HuGene-1.0-st-v1 and HG-U133 Plus 2) on airway epithelial cell samples collected from nasal (GSE19190) or bronchial (GSE11906) brushings of healthy non-smokers were obtained from the NCBI Gene Expression Omnibus (GEO) database.[128,312] This resulted in a total of 80 individual samples from the two different experiments that included 11 upper airway samples (nasal: $n = 11$) and 69 lower airway samples (trachea: $n = 17$, large airway: $n = 17$, small airway: $n = 35$). For all data set samples, raw intensity values and annotation data were downloaded using the GEOquery R package (version 2.52.0)[99] from the Bioconductor project.[174] Probe definition files were downloaded from Bioconductor and probes were annotated using Bioconductor's annotate package. All gene expression data were unified into a single data set that was then normalized by robust multiarray average (RMA) normalization, and only genes present in both of the Affymetrix platforms ($n = 16013$) were kept for subsequent analyses. Correction of experiment-specific batch effects was performed using the ComBat method[195] implemented using the sva R package (version 3.32.1).[229] RMA-normalized expression levels for conventional (*ACE2*, *TMPRSS2*, *ADAM17* and *CTSL* (cathepsin L1)) and non-conventional (*CD147* and *GRP78*) SARS-CoV-2 receptor genes were compared across the four defined airway levels, with *CDH1* (E-cadherin) expression level included as a positive control with known expression in lung tissue. Gene expression levels were tested for significant differences via pairwise Wilcoxon rank sum tests with Benjamini–Hochberg multiple testing correction using the stats R package (version 3.6.1). Gene expression box plots were generated with the ggplot2 R package (version 3.2.1).

Analysis of curated bronchial epithelial cell brushing data set

A total of 1859 public microarray experiments using Affymetrix chips (HG-U133 Plus 2 and HuGene-1.0-st-v1) on airway epithelial cell samples were selected from the NCBI GEO database. These samples were further filtered by removing individuals with asthma or COPD, resulting in a total of 504 individual healthy samples (GSE4302, 28 samples; GSE67472, 43 samples; GSE37147, 159 samples; GSE108134, 274 samples). Within this data set, sex and/or age information was included for 310 samples; of these, sex data were available for 86 females and 106 males. Smoking status information was also provided for 451 samples, with 260 current smokers, 82 former smokers and 109 never-smokers. For all data set samples, raw intensity values and annotation data were downloaded as described above. Probe definition files were retrieved as described above. All gene expression data were unified into a single data set that was then RMA-normalized, and only genes present in

both of the Affymetrix platforms ($n = 16105$) were kept for subsequent analyses. Correction of experiment-specific batch effects was performed as described above.

Analysis of promoter activity from the FANTOM5 data set

Promoter activity analysis performed by Dr. Michael Mansfield at the Okinawa Institute of Science and Technology, Japan

The FANTOM5 promoterome data set[91] for the hg38 assembly[7] was used to examine promoter activity of SARS-CoV-2-related human genes, namely ACE2, TMPRSS2, ADAM17, CTSL, CD147 and GRP78. Using the ZENBU genome browser,[340] the nearest cap analysis of gene expression (CAGE) peak upstream and on the same strand as each of the aforementioned genes was extracted and analysed. The data set consists of CAGE promoter activity data for 1866 primary cells, cell lines and tissues from humans, and is quantified as normalized transcripts per million (TPM). A subset of FANTOM5 CAGE data (120 samples) is presented, considering only samples related to lung, gut, heart and prostate tissues (consisting of 74, 19, 15 and 12 samples, respectively). Normalized TPM values for each CAGE peak, an approximation for promoter activity, were $\log_{10}$ transformed and separated according to tissue and cell type, and the radius of each point is proportional to these transformed normalized TPM values.

Analysis of protein abundance from proteomic data sets

Publicly available human proteomic data from the data sets of Kim *et al.*[211] and Foster *et al.*[119] were used to evaluate SARS-CoV-2 receptor-related protein expression in different human tissues and experimental conditions. Expression values were extracted from the data set of Kim *et al.*[211] for ACE2, TMPRSS2, ADAM17, CTSL, CD147 and GRP78, using CDH1 as a control for airway cells. Data were created using the pheatmap package in R (version 1.0.12) and expressed as $\log_{10}$-transformed to facilitate visualization. Proteomic data from the data set of Foster *et al.*[119] consist of bronchial epithelial cells collected from healthy non-smokers ($n = 4$; males) and exposed to PBS control vehicle. Intensity values for ACE2, TMPRSS2, ADAM17, CD147 and GRP78 were extracted, with CDH1 included as a positive control. Intensity values were determined by the original study authors via normalization of all detected peptide intensities associated with a given parent protein.[119] Box plots were generated with the ggplot2 R package (version 3.2.1) with intensity values $\log_{10}$-transformed for visualization purposes.

Analysis of single cell RNA sequencing data

scRNA-seq analysis performed by Dr. Anna Dvorkin-Gheva at McMaster University

Data pre-processed using the Cell Ranger pipeline (10x Genomics) were obtained from GSE135893. Samples from 10 control subjects and 12 idiopathic pulmonary fibrosis patients were downloaded and post-processed with the Seurat package in R.[64] Cell populations were defined using the markers provided in the source paper.[145] Cells belonging to the 10 control subjects were used for further analysis. Visualizations of violin plots were created using Seurat.

Primary human airway epithelial cells

Cell culturing performed by Abiram Chandiramohan, Nicholas Tiessen, and Quynh Cao at McMaster University

The human lung adenocarcinoma cell line, Calu-3, was grown under culture conditions defined by the supplier (ATCC HTB-55). Primary human airway epithelial cells isolated via bronchial brushings from consented healthy individuals were grown in PneumaCult ExPlus (Stemcell Technologies, Vancouver, BC, Canada) under submerged monolayer culture conditions and used between passages 1 and 4. The human bronchial epithelial cell line, HBEC-6KT, was grown under submerged monolayer culture conditions in keratinocyte serum-free media supplemented with epidermal growth factor (0.4 ng x mL^{-1}) and bovine pituitary extract (50 µg x mL^{-1}).[162,165,181,313]

Immunoblots

Immunoblots performed by Abiram Chandiramohan, Nicholas Tiessen, Quynh Cao, Louise Organ, Chitra Joseph, and Alison John at McMaster University

Cell protein was isolated using RIPA lysis buffer (VWR, Mississauga, ON, Canada) supplemented with protease inhibitor cocktail (Sigma, Oakville, ON, Canada) with quantification performed using Bradford assay reagents (Bio-Rad, Mississauga, ON, Canada). Immunoblots were performed using stain-free 4–20% pre-cast gradient gels and imaged on a ChemiDoc XRS+ Imaging system (Bio-Rad). For each immunoblot, 20 µg of protein was added per lane. ACE2 (MAB933, monoclonal, clone 171606, 2 µg x mL^{-1} (R&D Systems)), TMPRSS2 (HPA035787, polyclonal, 0.4 µg x $^{-1}$ (Atlas Antibodies)), CD147 (ab666, monoclonal, clone MEM-M6/1, 1 µg x $^{-1}$ (Abcam)) and GRP78 (610979, monoclonal, clone 40/BiP, 0.25 µg x mL^{-1} (BD Biosciences), and HPA038845, rabbit polyclonal (Atlas

Antibodies)) primary antibodies were diluted in 5% skimmed milk/TBS with 0.1% Tween-20 and incubated overnight on a rocker at 4°C with detection performed the following day using an anti-mouse (ACE2, CD147 and GRP78 (BD Biosciences)) or anti-rabbit (TMPRSS2 and GRP78 (Atlas Antibodies)) horseradish peroxidase (HRP)-conjugated secondary antibodies at 1:3000 for 2 h at room temperature (Cell Signaling, Danvers, MA, USA). visualization of TMPRSS2, CD147 and GRP78 was performed using Clarity Western enhanced chemiluminescence (ECL) Substrate, while ACE2 was visualized with Clarity Max ECL Substrate (Bio-Rad). Total protein loading images were collected as a qualitative visualization of protein loading between sample types.[144] The immunogen for ACE2 primary antibody is mouse myeloma cell line NS0-derived recombinant human ACE2 Gln18-Ser740 (predicted). The immunogen for TMPRSS2 primary antibody is the recombinant protein epitope signature tag antigen sequence, GSPPAIGPYYENHGYQPENPYPAQPTVVPTVYEVHPAQYYPSPVPQYAPRVLTQASNPVVCTQPKSPSGTVCTSKT.

The immunogen for the CD147 primary antibody is recombinant full-length protein corresponding to human CD147. The immunogen for the GRP78 BD Biosciences primary antibody is human BiP/GRP78 amino acids 525–628. The immunogen for the GRP78 Atlas Antibodies primary antibody is the recombinant protein epitope signature tag antigen sequence, EKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDFKAKKKELEEIVQPIISK.

Independent immunoblot analysis (L. Organ, C. Joseph, A. John and G. Jenkins) was performed on A549, HEK and immortalized human bronchial epithelial cells. Equal amounts of protein (20 µg) were loaded on to 4–12%, Bis-Tris gradient gels (NP0326BOX; ThermoFisher) with anti-ACE2 (ab108252, rabbit monoclonal, clone EPR4435(2), 1/500 dilution of stock antibody; Abcam). A loading control of GAPDH was used to demonstrate protein loading (ab181603, rabbit monoclonal, EPR16884, 1/10000 dilution of stock antibody; Abcam). Visualization was performed with ECL Clarity (Bio-Rad) on a Licor C-DiGit.

Immunohistochemistry

Immunohistochemistry performed by Spencer Revill at McMaster University

Formalin-fixed paraffin-embedded human lung tissue from non-diseased regions was obtained from archived tissue blocks from patients who had undergone lung resection for clinical care. Human heart tissue was from the University of British Columbia Cardiovascular Tissue Registry. Sections 4 µm thick were cut and stained for ACE2 (15 µg x mL^{-1}), TMPRSS2 (10 µg x mL^{-1}), CD147 (5 µg x mL^{-1}), and GRP78 (HPA038845, 1/200 dilution) using the same antibodies used for immunoblot analysis. All staining was performed on

a Leica Bond RX system with Leica Bond reagents, heat-induced antigen retrieval at pH 6 (20 min) with primary antibody incubation for 20 min. Digital slide scanning was performed using an Olympus VS120-L100 Virtual Slide System at 40× magnification with VS-ASW-L100 V2.9 software and a VC50 colour camera, followed by image visualization with HALO image analysis software.

4.1.3 Results

Candidate genes important in SARS-CoV-2 infection are detectable at varying levels in human airway epithelial cells and lung tissue

I performed a targeted analysis of *ACE2*, *TMPRSS2*, *ADAM17*, *CTSL*, *CD147* and *GRP78* gene expression as candidates important for SARS-CoV-2 infection in human airway epithelial cells. Here and throughout the gene expression analyses, *CDH1* (E-cadherin) was used as a control for lung epithelial cell phenotype. I first examined these genes in a curated data set of upper and lower airway epithelial cell gene expression from the nasal sinus to the 12th generation of airway in the lung (Figure 4.1).

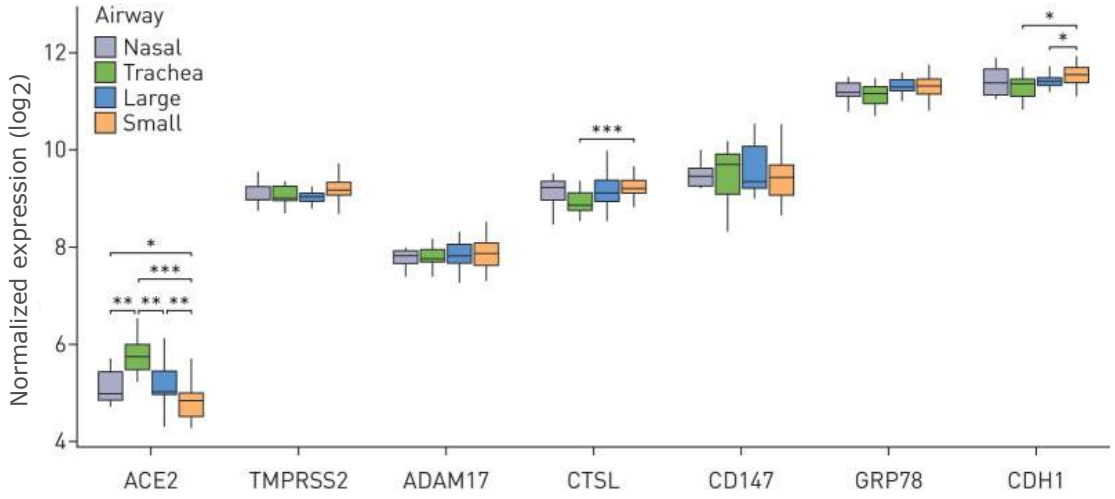


Figure 4.1: Normalized $\log_2$ expression levels for *ACE2*, *TMPRSS2*, *ADAM17*, *CTSL*, *CD147* and *GRP78* genes compared across the upper airway (nasal; $n = 11$) and lower airways (trachea; $n = 17$, large airway; $n = 17$, and small airway; $n = 35$). *CDH1* (E-cadherin) gene expression level is included as a positive control. Statistical values for comparisons for each gene at each airway generation were calculated; those not shown were non-significant. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

In the upper airways, all candidates were expressed, with the highest level observed for *GRP78* and the lowest level observed for *ACE2*. Analysis along multiple generations of the lower airways (trachea, large (4th–6th generation) and small airways (10th–12th generation)) revealed identical relative expression patterns, with *ACE2* being the least expressed and *GRP78* being the highest expressed. *ACE2* gene expression showed the greatest variability along the upper and lower airways, with greatest expression observed in the trachea samples and the lowest expression in the small airway (Figure 4.1).

Following the observation of consistent expression along the upper and lower airways of candidate genes important in SARS-CoV-2 infection, I determined whether sex or age affected gene expression levels in healthy individuals using a curated data set of bronchial brushings from 504 healthy subjects (Table 10). The expression levels for the candidate genes in healthy subjects paralleled the patterns observed in the smaller survey of upper airways, trachea, large and small airways (Figure 4.2). Median *ACE2* gene expression was the lowest, while *GRP78* gene expression was the highest (Figure 4.2 A). No gene candidate demonstrated sex dependence for expression levels (Figure 4.2 B). *ACE2*, *TMPRSS2*, *CD147* and *GRP78* were elevated in current smokers relative to never-smokers (Figure 4.2 C and Table 11). *CTSL* was reduced in current smokers relative to never-smokers. No microarray chip-dependent effects were observed for relationships between sex or smoking status and gene expression. For quantitative analyses related to age and gene expression, the curated database was divided into data sets that used either the HG-U133 Plus 2 or HuGene-1.0-st-v1 microarray due to differences in age distributions. In the HuGene-1.0-st-v1 data set ($n = 181$), which included a greater proportion of older individuals (>50 years), reduced *ACE2* gene expression was observed with age (Figure 4.2 E; $p < 0.05$).

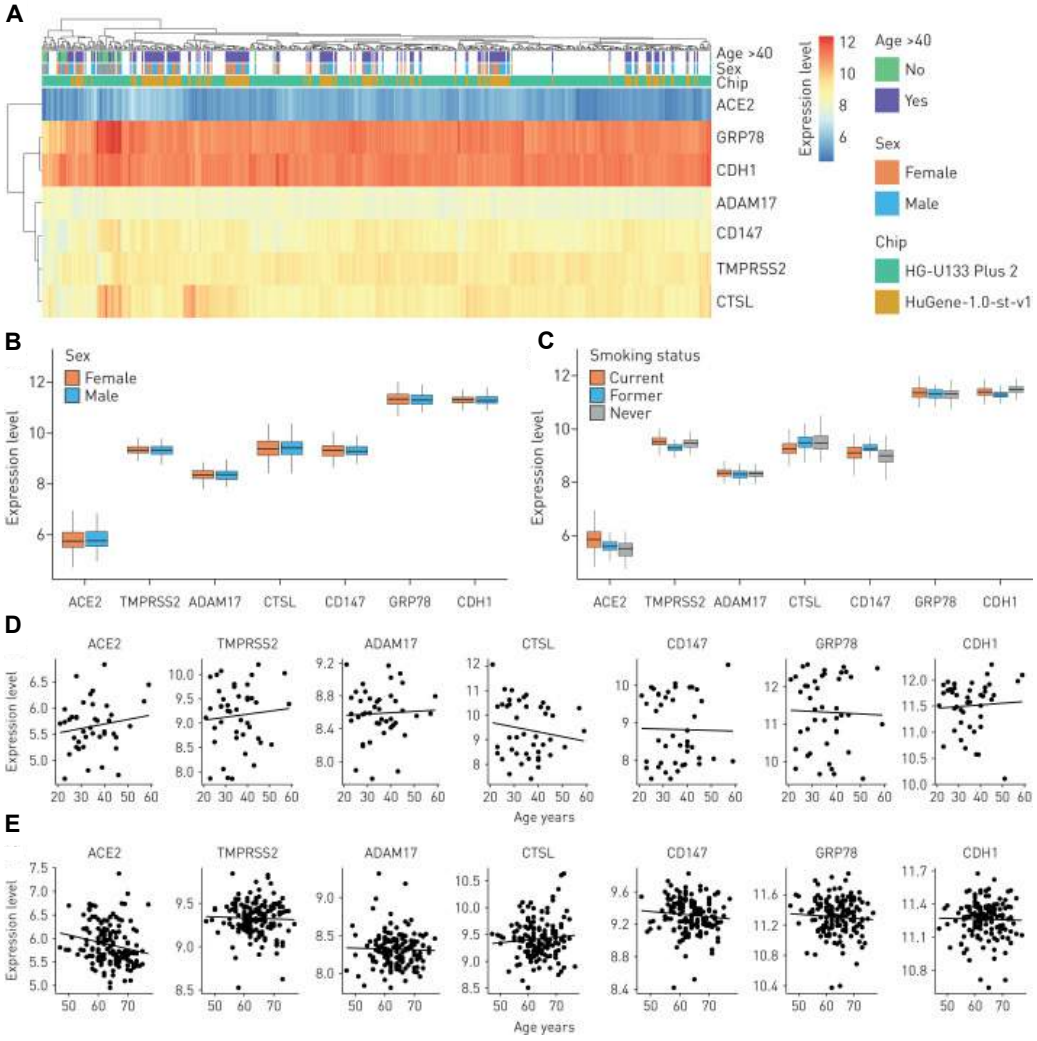


Figure 4.2: (A) Clustered heat map of $\log_2$ expression levels from NCBI Gene Expression Omnibus (GEO) samples ($n = 504$), annotated by age, sex and microarray chip platform. Expression values reflect signal intensities, indicating lowest detected expression of *ACE2* and highest expression of *GRP78* and *CDH1*. (B) and (C) Box plots of expression levels separated by (B) sex ($n = 194$) and (C) smoking status ($n = 451$). (D) and (E) Plots of gene expression levels versus age, with linear regression lines of best fit, for data sets that used either (D) the HG-U133 Plus 2 microarray ($n = 43$) or (E) the HuGene-1.0-st-v1 microarray ($n = 181$). Correlations were performed separately between platforms because of differences in their age distributions. (E) A weak negative correlation ($r = -0.20$, $p = 0.015$) was detected for *ACE2* in the data set that used the HuGene-1.0-st-v1 microarray.

Promoter activity data of each of the candidate genes important in SARS-CoV-2 binding and infection were extracted and analysed from the FANTOM5 data set, which includes 1866 primary cells, cell lines, and tissue sample types (Figure 4.3). All sample formats that included "lung", "nasal", "airway" or "olfactory" were selected, to identify lung-specific sample types. Gut, heart and prostate tissue samples were analysed as controls. Consistent with the observed gene expression analysis along the upper and lower airways, normalized TPM values for each CAGE peak demonstrated that CD147 promoter activity was elevated relative to *ACE2* promoter activity across airway epithelial cells and lung tissue samples. *CTSL* promoter activity was the lowest of all candidate genes, which contrasted with the modest expression observed at the gene level (Figure 4.2 A). Both microarray gene expression analysis and promoter activity were consistent with results of candidate gene expression in a scRNA-seq data set of 10 healthy subjects (Figure 3).

Collectively, this gene expression analysis of the upper and lower airways of healthy males and females of diverse ages suggests that *ACE2* gene expression is low relative to all other candidate SARS-CoV-2 receptor genes analysed in human airway epithelial cells. Furthermore, no sex-dependent or age-dependent expression patterns were observed for any candidates at the gene level, although smoking status did have an impact on gene expression levels.

In vitro and *in situ* protein profiling reveals distinct expression patterns for candidates important in SARS-CoV-2 infection

Analysis of transcriptional data may not be indicative of *in situ* protein expression levels.[400] To extend these gene expression observations, I mined publicly available proteomic data from whole lung and primary human airway epithelial cell cultures and performed *in vitro* immunoblots on human airway epithelial cell lysates and *in situ* protein immunohistochemistry on human lung tissue using the same antibodies for each method.

The Human Proteome Map is a publicly available resource that includes select adult and fetal tissues and circulating immune cell populations.[211] Using this resource, I examined protein expression of ACE2, TMPRSS2, ADAM17, CTSL, CD147 and GRP78. In human lung tissue homogenate, ACE2 was not detected, while being detected in heart, gut and testes, known positive control tissues (Figure 4.4 A). The rank order of the remaining molecules in human lung tissue homogenate was: GRP78 > CD147 > CTSL > ADAM17 > TMPRSS2. Human lung tissue homogenate is a heterogeneous population of cells, precluding the ability to associate protein expression to a given cell type. I therefore interrogated a publicly available proteomic data set derived from primary human airway epithelial cells grown under air–liquid interface culture conditions,[119] examining the same candidates.

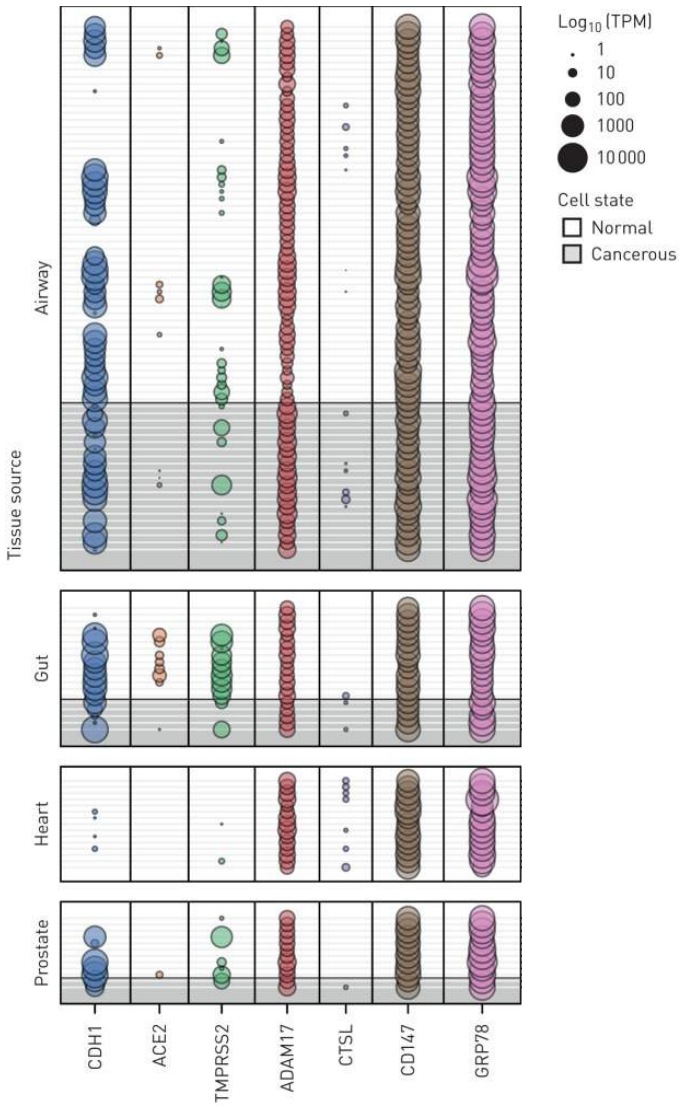


Figure 4.3: The FANTOM5 CAGE data consist of quantified promoter expression levels across the human genome for 1866 samples from primary cells, cell lines and tissue samples. The FANTOM5 CAGE promoter activity data for several SARS-CoV-2-related genes are shown for samples related to lung, gut, heart and prostate tissues ($n = 120$). Dot sizes are proportional to promoter activity, depicted as $\log_{10}$-transformed normalized transcripts per million (TPM). Notably, *ACE2* is either not expressed or expressed at low levels (<1 TPM in all but one sample) in the airway, including measurements from healthy and cancerous cells.

Again, ACE2 protein expression was not detectable (Figure 4.4 B). CD147, GRP78 and CTSL were expressed with multiple peptide counts, while TMPRSS2 and ADAM17 were only marginally expressed with low peptide counts. Collectively, two proteomic data sets from distinct lung sample formats provide complementary and consistent expression profiles of candidate molecules important in SARS-CoV-2 infection.

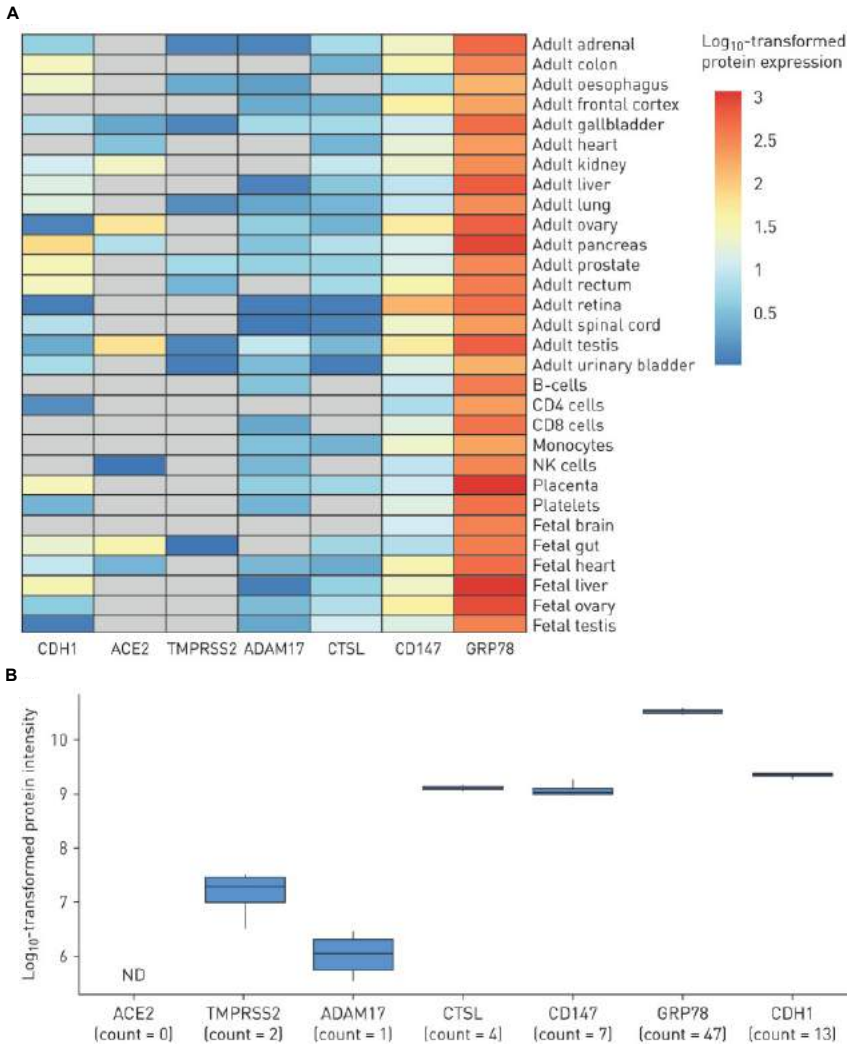


Figure 4.4: (A) Intensity values of protein expression from Kim *et al.*[210] for the genes ACE2, TMPRSS2, ADAM17, CTSL, CD147 and GRP78. CDH1 intensity is included as a positive control for expression in airway cells. Intensity values have been $\log_{10}$-transformed to facilitate comparison between candidates with different basal expression levels across tissue types. Grey cells in the heat map correspond to an untransformed intensity of 0 and represent an undetectable signal. (B) Intensity values $\log_{10}$-transformed for visualization of ACE2, TMPRSS2, ADAM17, CTSL, CD147 and GRP78 proteins in human airway epithelial cells from healthy non-smokers ($n = 4$; males) grown under air–liquid interface culture conditions.[119] CDH1 intensity is included as a positive control. Counts indicating the number of detected peptides associated with each parent protein are provided. ND: the protein was not detected in this study.

To localize the *in situ* expression of the candidate molecules of interest at the protein level, collaborators pursued immunohistochemical analysis paired with immunoblot validation of the specificity of the selected antibodies for recognition of proteins of the predicted molecular weight. An anti-ACE2 antibody detected only a single band in Calu-3 cells at the predicted molecular weight of ACE2 protein ($\sim$ 110 kDa) (Figure 4.5 A, lanes 1–3). The anti-ACE2 antibody required the use of a super-sensitive ECL solution. No ACE2 protein was detected in primary airway epithelial cells or the HBEC-6KT cell line, despite confirmation of protein loading (Figure 4.5 A, lanes 4–9, protein loading shown underneath main blot). Independent immunoblotting with a distinct anti-ACE2 primary antibody was performed, with a single band observed in HEK cells, but not in immortalized human bronchial epithelial cells or A549 cells (Figure 4).

An anti-TMPRSS2 antibody detected multiple bands in all airway epithelial cell samples, with a dominant band at the predicted molecular weight of 57 kDa (Figure 4.5 B). These patterns were conserved across all cell types that were analysed. An anti-CD147 antibody detected a single band in all airway epithelial cell samples, with a dominant band at the predicted molecular weight of 55 kDa (Figure 4.5 C). The immunoblot bands were consistent with the heavy glycosylation of CD147.[196] An anti-GRP78 antibody (BD Biosciences 610979) detected a single band in all airway epithelial cell samples, with a dominant band at the predicted molecular weight of 78 kDa (Figure 4.5 D).

The immunoblots using anti-ACE2, anti-CD147 and anti-GRP78 demonstrated a single band of predicted molecular weight, suggesting that observed immunohistochemical staining should be specific to the protein of interest based on the target epitope, as both methods detect denatured proteins.[389] The same anti-ACE2 and anti-CD147 antibodies were validated for immunohistochemistry. The anti-TMPRSS2 was used for immunohistochemistry, although the multiple bands observed by immunoblot caution the specificity of any observed *in situ* staining. Attempts to optimise anti-GRP78 antibody application for immunohistochemistry were unsuccessful, requiring additional antibody interrogation with HPA038845 (Atlas Antibodies), which was suitable for both immunoblotting and immunohistochemistry (Figure 5).

ACE2 immunohistochemistry revealed only select staining in rare cells in the airways and the alveoli of all 98 human lung samples analysed, which included healthy subjects and those with chronic lung diseases (Figure 4.6). A single healthy human sample contained one positive airway epithelial cell with additional positive staining in the peripheral lung in cells with type II alveolar epithelial cell morphology (Figure 4.6 A). A representative image of a sample from a smoker with COPD (Figure 4.6 B) shows no ACE2 protein staining in the airway epithelium and a rare positive cell in sub-basement membrane tissue. Quantification of positive pixel count for ACE2 staining normalized to total tissue pixel

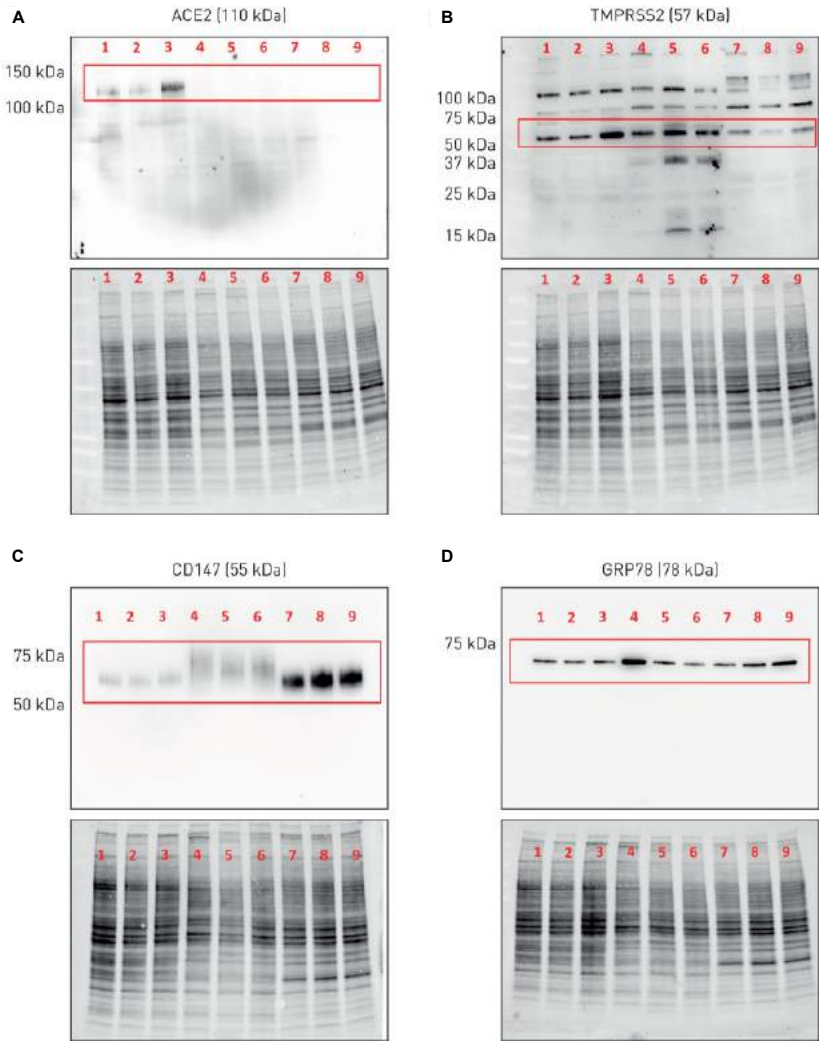


Figure 4.5: (A) ACE2 with single band for predicted molecular weight of 110 kDa (red box). (B) TMPRSS2 with multiple bands including a dominant band at predicted molecular weight of 57 kDa (red box). (C) CD147 with a single broad band around predicted molecular weight of 55 kDa (red box). (D) GRP78 with a single band at predicted molecular weight of 78 kDa (red box). Lanes 1–3: Calu-3 cells. Lanes 4–6: primary human airway epithelial cells. Lanes 7–9: human bronchial epithelium cell (HBEC)-6KT cell line. All cells grown under submerged monolayer conditions, with $n = 3$ independent passages (Calu-3 or HBEC-6KT) or donor samples (primary human airway epithelial cells; non-smoker, healthy subjects). For each independent blot of each protein, all of the same samples were run. A total protein loading control is provided to demonstrate protein loaded for each sample.

count revealed no differences between healthy non-smokers and tobacco smokers (Figure 6). Lung microvasculature and human heart tissue had positive staining (Figures 7 and 8), consistent with previously described reports for ACE2 protein staining patterns.[63,106]

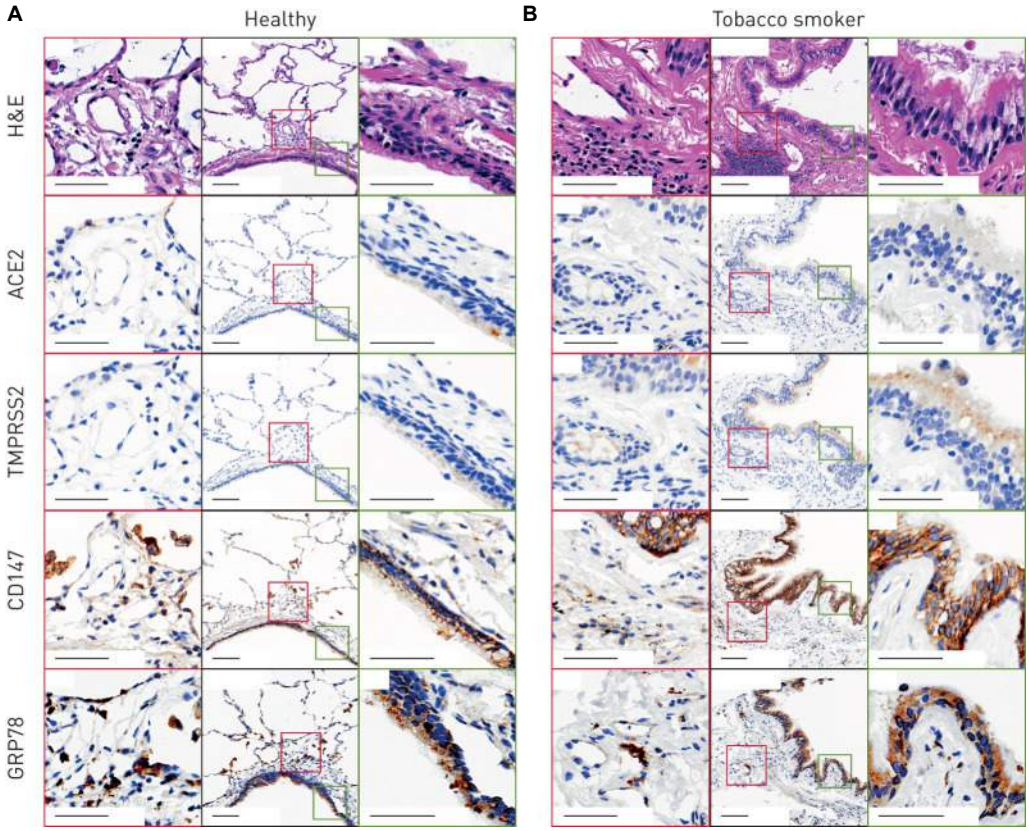


Figure 4.6: Representative samples from (A) a healthy non-smoker with no underlying chronic airway disease, and (B) a smoker with COPD. Black outlines: low magnification (12x) of conducting airways with airway epithelium; scale bars 100 µm. Green outlines: high magnification regions (60x) of conducting airway epithelium that are defined in the low magnification images by green squares; scale bars 50 µm. Red outlines: high magnification regions (50x) of lung tissue away from the airway lumen that are defined in the low magnification images by red squares; scale bars 50 µm. H&E: haematoxylin and eosin. Positive immunohistochemical staining is rust/brown. Total number of independent samples analysed was 49–98.

TMPRSS2 immunohistochemistry revealed diffuse staining in the airway epithelium

and in immune cells in the lung periphery, with greater staining in smokers with COPD (Figure 4.6). These observations were consistent in the majority of the 98 human samples examined. CD147 immunohistochemistry revealed strong membrane-restricted staining in the airway epithelium and diffuse staining in immune cells in the lung periphery (Figure 4.6). CD147 displayed greater staining in smokers with COPD. These observations were consistent in the majority of the 98 human samples examined.

GRP78 immunohistochemistry revealed diffuse staining in airway and alveolar epithelium and in immune cells in the lung periphery (Figure 4.6; 49 samples). No qualitative differences in GRP78 staining were observed between healthy subjects and smokers with COPD. TMPRSS2 and CD147 protein expression are potentiated in individuals with a history of tobacco smoking and a diagnosis of COPD.

Collectively, the *in vitro* and *in situ* protein profiling in this study is consistent with the gene expression analysis, with CD147 and GRP78 protein expression dominant over TMPRSS2 and ACE2. Additional examples of staining in lung tissue are provided in Figure 9. ACE2 protein expression is rare in human lung tissue and found in select cells in both healthy individuals and those with chronic lung diseases.

4.1.4 Discussion

The COVID-19 pandemic that emerged in late 2019 is caused by SARS-CoV-2. The possible host receptor(s) for SARS-CoV-2 have not been exhaustively surveyed in human lung tissue at the gene and protein level. Understanding the expression levels and localization of candidate SARS-CoV-2 receptors in host lung tissue may provide insights into therapeutic interventions that might reduce disease spread, viral replication or disease pathology. To address this knowledge gap, gene expression, proteomic profiling at the tissue and cell level, and *in situ* protein profiling of candidate receptors in human airway epithelial cells and lung tissue were performed (summarized in Figure 4.7). Collectively, the data demonstrate rare ACE2 protein expression in human airway epithelial cells *in vitro* and *in situ*. The protein expression data are consistent with low ACE2 promoter activity in a panel of lung epithelial cell samples and low *ACE2* gene expression in bronchial epithelial cells (microarray) and lung cells (scRNA-seq). This study presents confirmatory evidence for the presence of TMPRSS2, CD147 and GRP78 protein *in vitro* in airway epithelial cells and confirm broad *in situ* protein expression of CD147 and GRP78 in the respiratory mucosa. The data suggest that for ACE2 to be an integral receptor for SARS-CoV-2, mechanisms are likely to exist that dynamically regulate expression in human lung, perhaps in periods of SARS-CoV-2 infection.[444] It is also possible that alternate receptors for SARS-CoV-2 are important in initial host cell infection.

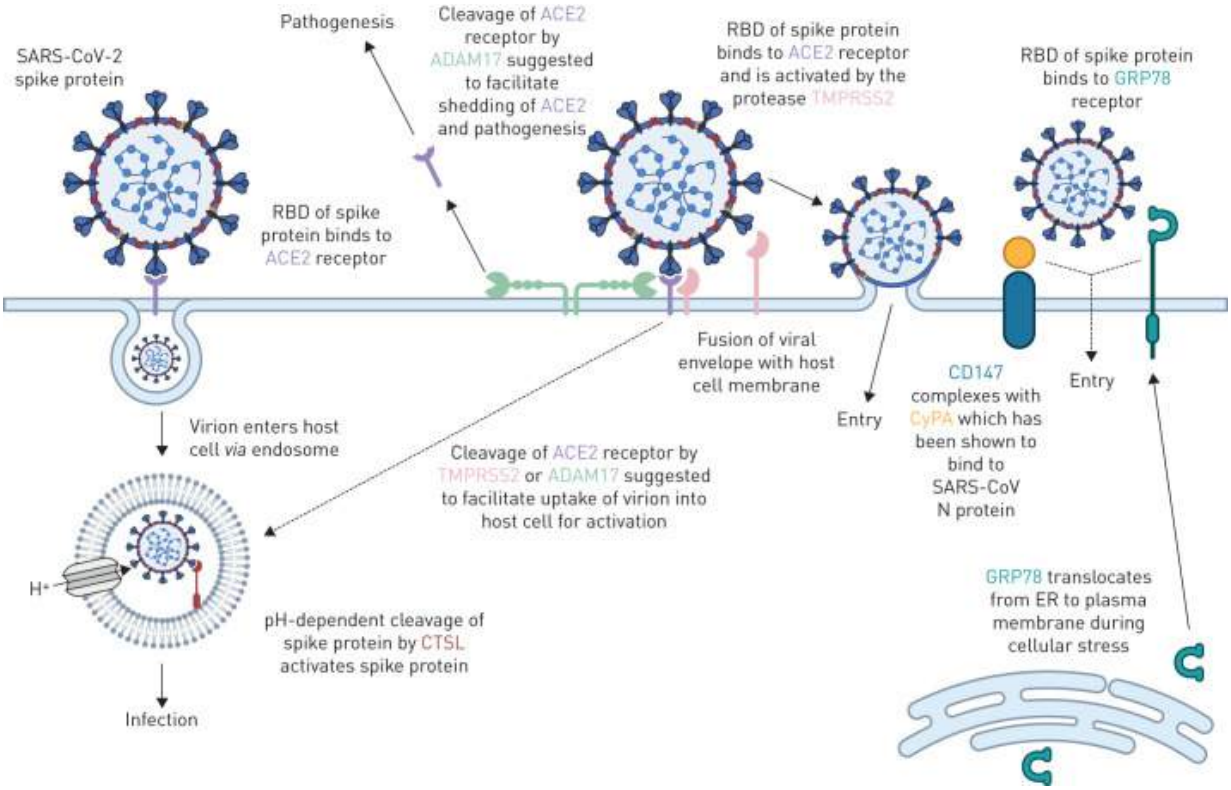


Figure 4.7: Proteins associated (or suggested to be associated) with host cell entry of SARS-CoV-2 and the activation of the SARS-CoV-2 spike protein (SARS-S) are displayed. Angiotensin-converting enzyme 2 (ACE2) is suggested as the primary SARS-S receptor for viral entry (interaction of ACE2 with SARS-S receptor-binding domain (RBD) leading to endosomal viral uptake), followed by activation of SARS-S via pH-dependent cleavage mediated by cathepsin L1 (CTSL). Secondary methods of viral entry and SARS-S activation are likely to involve proteases (e.g., TMPRSS2 and ADAM17) and/or secondary receptors (CD147 and GRP78). Dashed lines indicate mechanisms that have not been fully validated. CyPA: cyclophilin A; N protein: nucleocapsid protein; ER: endoplasmic reticulum. Adapted from Heurich *et al.*[158] with updates and additional information on candidate host molecules. Figure created with BioRender.com.

Using a curated microarray gene expression data set generated from bronchial brushings of 504 healthy subjects that considers the limitations of merging multiple data sets from distinct sources, it was observed that sex did not correlate with gene expression of any candidate host molecule involved in SARS-CoV-2 infection and that *ACE2* and *TMPRSS2*

were the lowest expressed genes of interest examined. In one data set, *ACE2* gene expression modestly decreased with age, although protein level confirmation was not possible. The low level of *ACE2* and *TMPRSS2* gene expression in bulk bronchial epithelial cell gene expression samples suggests low levels of cells expressing both of these genes within this lung tissue. This study confirms that tobacco smoking is associated with elevated *ACE2* gene expression levels in bronchial epithelial cell samples,[233] although this was unable to be confirmed with immunohistochemistry analysis of protein on human lung samples.

Advances in transcriptomics have enabled scRNA-seq that has identified unique and rare cell types in human lung that may have importance in health and disease.[305,336] scRNA-seq provides an opportunity to look at transcriptional profiles in subsets of cell populations, which may isolate a cell signal from a bulk sample. Therefore scRNA-seq data from healthy human lung samples was utilized as a parallel approach. The resolution of scRNA-seq for sub-populations of epithelial cells revealed low or absent expression of *ACE2* gene in all populations examined, whereas *CD147* and *GRP78* were present in all populations. These results are consistent with current publicly available data that discuss the presence of rare ACE2/ TMPRSS2-positive cells.[444] Using lung samples from eight individuals (four HIV and active tuberculosis double positive, two HIV positive and tuberculosis negative, and two double negative controls), Ziegler *et al.*[444] have reported in humans that only 0.8% of type II alveolar epithelial cells expressed both *ACE2* and *TMPRSS2* genes. Further analysis of ciliated cells found that 5.3% of these cells expressed both *ACE2* and *TMPRSS2* genes. *in vitro* models with SARS-CoV are consistent with this finding, as ciliated cells are preferentially targeted by this coronavirus.[351] Most intriguing is that *ACE2* and *TMPRSS2* gene-expressing cells were only identified in the HIV and tuberculosis double positive samples. These observations were replicated in the upper airways, with only a rare population of secretory epithelial cells (0.3% of this population) co-expressing *ACE2* and *TMPRSS2*. The reported scRNA-seq results are consistent with a focused analysis looking at only *ACE2* gene expression in a variety of lung cell types.[311] Importantly, these elegant transcriptomic analyses confirm the observations in bulk tissue microarray data sets.

Consortium-based publicly available data sets represent another parallel approach to confirm the data. This study used the FANTOM5 data set containing CAGE promoter activation data for 1866 primary cells, cell lines and tissue samples from humans[91] to examine the level of promoter activity for each candidate SARS-CoV-2 receptor gene. The FANTOM5 CAGE data provide an additional and complementary approach to quantifying gene expression, since a given gene's shared promoter can yield multiple transcripts at different expression levels, as well as being partially independent of any given transcript's half-life in the cell. In general, the promoter activity of *ACE2* in airway-related tissues is low or absent; only a single sample originating from an adult lung yielded a normalized

CAGE promoter expression level >1 TPM, while expression was observed in gut cells, consistent with known patterns of *ACE2* expression.[161] Consistent with the microarray data, *CD147* promoter activity is elevated relative to *ACE2* across airway-related cells and tissues, although the relatively low *CTSL* promoter activity is incongruent with modest levels of gene expression.

The expression of genes does not always correlate with protein expression.[400] With this in mind, combination proteomic analyses with immunoblot analyses were performed. For the immunoblots, the human Calu-3 adenocarcinoma cell line was used, as this cell line is susceptible and permissive to SARS-CoV-2 infection and expresses *ACE2*, an observation this study confirms.[169,385] This study also used primary human airway epithelial cells and the bronchial epithelial cell line (HBEC-6KT). Immunoblots for ACE2 and TMPRSS2 were performed, as these have been highlighted as interacting with SARS-CoV-2, and CD147 was probed as recent pre-clinical and clinical studies have provided proof of concept for this as a candidate SARS-CoV-2 receptor.[49,407] Lastly, *GRP78* was dominantly expressed throughout transcriptomic studies and was selected as a positive control, as previous expression has been confirmed in human airway epithelial cells.[15] CTSL was excluded from the present analysis due to low promoter activity (Figure 4.3), while ADAM17 was excluded as the proposed function in coronavirus infections is via ACE2,[169,440] which was included in analysis. Immunoblot analysis with all antibodies revealed dominant bands of predicted molecular weight, with the anti-TMPRSS2 polyclonal antibody revealing additional minor bands in all cell samples examined. The identity of these other bands remains unclear and suggests downstream immunohistochemical analysis may be confounded by the specificity of this antibody. In contrast, antibodies for ACE2, CD147 and GRP78 were specific and could be used for immunohistochemistry without concerns of specificity. Interestingly, ACE2 protein could only be detected with a super-sensitive ECL solution and only in Calu-3 cells, suggesting absent protein in primary human airway epithelial cells and the HBEC-6KT cell line. The data in this study are consistent with previous immunoblots of primary human airway epithelial cells grown under submerged monolayer conditions using the same primary antibody, where ACE2 protein was absent, and only expressed under air–liquid interface culture conditions.[194] The observation that CD147 and GRP78 are also expressed in Calu-3 cells encourages further interrogation into these host proteins, as they may contribute to function of ACE2 and TMPRSS2 in SARS-CoV-2 binding and fusion in this cell type. Collectively, the profiling of antibodies by immunoblot of airway epithelial cells revealed distinct band patterns demonstrative of antibody specificity for ACE2, CD147 and GRP78, and to a lesser extent for TMPRSS2.

Immunohistochemical analysis has been performed for localization of ACE2 and TMPRSS2 in human lung.[46,149] The observation of positive staining in human lung tissue

for these proteins was not accompanied by companion immunoblot or complementary approaches to define the specificity of the antibody used.[389] In the absence of determination of antibody specificity, the historical data presented should be interpreted with caution. To address the issue of antibody specificity for immunohistochemical staining, the same antibodies validated by immunoblot were used and confirmed findings using proteomics as an orthogonal, antibody independent, approach. This study again focused on ACE2 and TMPRSS2 as these are candidate proteins important for SARS-CoV-2 infection of host cells. The immunohistochemical staining patterns of ACE2 observed in this study were consistent with transcriptional profiling and immunoblots with only one out of 98 human samples demonstrating rare staining in the airway and alveolar epithelium. Positive ACE2 staining in heart tissues and areas of lung microvasculature suggest the staining protocol used was successful. These results directly contrast with those reported using antibodies that lacked validation for specificity.[46,149] TMPRSS2 was expressed more frequently across all samples examined, with variability in the airway epithelium associated with history of smoking and/or COPD status. In contrast, CD147 expression was observed in airway epithelium of all samples. Similar to TMPRSS2, elevated CD147 expression was associated with history of smoking and/or COPD status, consistent with previous reports.[196] The original GRP78 antibody selected for immunoblotting was not validated for immunohistochemistry. Therefore confirmatory GRP78 immunoblotting and immunohistochemistry were performed with an additional antibody (HPA038845) and provided demonstration of expression in human airway epithelial cells *in vitro* and *in situ*. Importantly, it is well established that GRP78 resides in the endoplasmic reticulum (ER) under normal physiological conditions, where it acts as an ER-resident molecular chaperone to facilitate correct protein folding. However, under conditions of ER stress, including viral infection, a portion of ER-resident GRP78 relocates to the cell surface, where it may act as a viral co-receptor.[135,437] The presence of cell surface GRP78 has been reported in atherosclerotic plaques,[95] prostate cancer[16] and kidney.[218] However, there are currently no commercially available anti-GRP78 antibodies that bind specifically to cell surface GRP78 and dual immunofluorescence is used to show co-localization of cell surface GRP78 with an established surface receptor.[95] This drawback precludes this study's ability to perform accurate cell surface GRP78 immunohistochemistry on lung tissue to interrogate this concept further in the context of SARS-CoV-2 receptors. Utilisation of prostate cancer patient-derived GRP78 auto-antibodies that are specific for cell surface GRP78[95] may be suitable on lung tissue for assessment of cell susceptibility to SARS-CoV-2 infection by GRP78.

This study has several limitations that have not already been addressed. The observation of differences in gene expression between upper and lower airways and along the airway tree were not corroborated at the protein level. It remains possible that entirely different

protein expression profiles for the candidate molecules examined exist in the upper airway, presenting a different environment for SARS-CoV-2 interaction with the respiratory mucosa. Nasal pharyngeal swabs are capable of detecting SARS-CoV-2 virus[408] and this anatomical region is probably important for subsequent infection in the lower airways.[171,445] Related to this potential temporality of effect, it is possible that SARS-CoV-2 induces the expression of receptors on host cells following infection.[444] This study is also limited by examining candidate molecules important in SARS-CoV-2 infection under basal conditions, in the absence of viral or environmental stimuli that may regulate gene transcription and protein translation.

SARS-CoV-2 infection and transmission has caused the COVID-19 pandemic. An understanding of the receptors used by SARS-CoV-2 for host cell infection and the parallel characterisation in human samples is required to inform development of intervention strategies aimed at mitigating COVID-19. The data demonstrate rare ACE2 protein expression in human airway epithelial cells *in vitro* and *in situ*, consistent with low *ACE2* promoter activity and *ACE2* gene expression in bronchial epithelial cells. This study presents confirmatory evidence for the presence of TMPRSS2, CD147 and GRP78 protein *in vitro* in airway epithelial cells and confirm broad *in situ* protein expression of CD147 and GRP78 in the respiratory mucosa. Due to the overwhelming evidence that the SARS virus interacts with ACE2, there are likely to be alternate mechanisms regulating ACE2 in the respiratory mucosa in the context of SARS-CoV-2 infection, and/or perhaps other co-receptors, beyond what is expressed under basal conditions at the protein level.

4.2 Experimental and natural evidence of SARS-CoV-2 infection-induced activation of type I interferon responses

Material in this section has been published as part of Banerjee *et al.* (2021).[34] The published manuscript is available here:

> Banerjee, A., El-Sayes, N., Budylowski, P., Jacob, R. A., Richard, D., Maan, H., Aguiar, J. A., Demian, W. L., Baid, K., D'Agostino, M. R., Ang, J. C., Murdza, T., Tremblay, B. J., Afkhami, S., Karimzadeh, M., Irving, A. T., Yip, L., Ostrowski, M., Hirota, J. A., Kozak, R., Capellini, T. D., Miller, M. S., Wang, B., Mubareka, S., McGeer, A. J., McArthur, A. G., Doxey, A.C.,

Mossman, K. Experimental and natural evidence of SARS-CoV-2-infection-induced activation of type I interferon responses. *iScience*, 24(5): 102477. 2021.[34] https://doi.org/10.1016/j.isci.2021.102477

4.2.1 Introduction

Innate antiviral responses, which include type I interferons (IFNs), are the first line of antiviral defense against an invading virus.[204] Cellular pattern recognition receptors (PRRs) recognize viral nucleic acids and activate key cellular kinases, such as inhibitor of nuclear factor kappa-β kinase subunit epsilon (IKKϵ) and TANK-binding kinase 1 (TBK1). These kinases phosphorylate and activate transcription factors such as interferon regulatory factor 3 (IRF3) to stimulate downstream production of type I/III IFNs.[215] Type I IFNs interact with interferon alpha/beta receptor (IFNAR) on cells to induce phosphorylation and activation of downstream mediators, such as signal transducer and activator of transcription 1 and 2 (STAT1 and STAT2), which leads to the production of antiviral interferon-stimulated genes (ISGs). Similarly, type III IFNs interact with their cognate receptors, IL-10R2 and IFNLR1, to activate STAT1 and STAT2, followed by the production of ISGs.[260]

Viruses encode proteins that can inhibit type I IFN production and signaling.[202,339] Emerging pathogenic human coronaviruses, such as SARS-CoV and Middle East respiratory syndrome (MERS)-CoV, have evolved multiple proteins that inhibit type I IFN responses in human cells.[84,101,246,355,429] Thus, to better understand SARS-CoV-2 pathogenesis, it is critical to identify the dynamic interaction of SARS-CoV-2 and the type I IFN response. Emerging data suggest that ectopic expression of at least 13 SARS-CoV-2 proteins, namely NSP1, NSP3, NSP6, NSP12, NSP13, NSP14, NSP15, M, ORF3a, ORF6, ORF7a, ORF7b, and ORF9b, can inhibit type I IFN responses in human cells.[136,230,411,422] However, limited studies have captured the dynamic interplay of viral-RNA-mediated up-regulation of type I IFN responses, followed by subsequent modulation of these responses by SARS-CoV-2 proteins as they accumulate in infected cells. Understanding the mechanisms of IFN modulation by SARS-CoV-2 proteins remains an area of intense research. In the meantime, intriguing observations about SARS-CoV-2 proteins have been reported by different groups. For example, SARS-CoV-2 NSP15 has been reported as an IFN-modulating protein by Gordon *et al.*,[136] but Lei *et al.*[230] were unable to identify NSP15 as an inhibitor of IFN promoter activation. In addition, both Gordon *et al.* and Jiang *et al.* identified ORF9b as a modulator of IFN responses,[136,230,411] but the study by Lei *et al.* did not identify ORF9b as a modulator.[230] Furthermore, infection with wild-type SARS-CoV-2 in Caco-2 cells activated phosphorylation of TBK1 and IRF3, along with mild induction of ISGs.[347] More recently, Yin *et al.* have demonstrated that wild-type SARS-CoV-2 induces a delayed type

I IFN response via melanoma differentiation-associated protein 5 (MDA5) recognition.[430] Thus, in-depth studies with clinical isolates of SARS-CoV-2 are required to confidently identify type I IFN responses that are generated in infected human cells and to determine the dynamic induction and modulation of type I IFN responses by wild-type virus infection.

Transcriptional data from *in vitro* and *in vivo* work have demonstrated the lack of induction of type I IFN responses following SARS-CoV-2 infection.[52] In contrast, emerging data from patients with mild and moderate cases of COVID-19 have demonstrated the presence of type I IFN.[146,384] Subsequently, recent studies have identified type I IFN responses in severe COVID-19 cases, which have been speculated to be associated with an exacerbated inflammatory response.[441] In addition, up-regulation of ISGs was also identified in a single-cell RNA sequencing study of peripheral blood mononuclear cells (PBMCs) from hospitalized COVID-19 patients.[414] Studies with patient samples are critical to understand the pathogenesis of SARS-CoV-2; however, the timing of sample collection, case definition of disease severity, and varying viral load can lead to different observations related to IFN responses. An early and controlled IFN response is preferable during virus infection. Excessive induction of type I IFN responses in COVID-19 patients is associated with higher levels of damaging inflammatory molecules.[247] Thus, it is critical to identify the extent to which SARS-CoV-2 can induce or inhibit human IFN responses using controlled mechanistic studies.

In this study, global early transcriptional responses have been identified that are initiated during infection of human airway epithelial (Calu-3) cells at 0, 1, 2, 3, 6, and 12 hours post incubation with a clinical isolate of SARS-CoV-2 from a COVID-19 patient in Toronto.[35] Data from this study demonstrate that SARS-CoV-2 infection induces the expression of type I IFNs, along with the expression of downstream ISGs. This study also identified an increasing trend for type I IFN expression (IFN-α2) in sera from moderate cases of COVID-19, relative to healthy individuals and severe cases of COVID-19. *In vitro* infection with SARS-CoV-2 induced phosphorylation of canonical transcription factors that are involved in the type I IFN response, such as IRF3, STAT1, and STAT2; exogenous activation of these transcription factors was not inhibited by wild-type SARS-CoV-2. In addition, higher serum levels of anti-inflammatory cytokines were detected in moderate cases of COVID-19 than in severe cases. Severe cases of COVID-19 displayed higher serum levels of pro-inflammatory cytokines. Data from this study suggest that replication-competent SARS-CoV-2 induces type I IFN responses in human airway epithelial cells, and type I IFN (IFN-α2) level detected in patients with moderate COVID-19 is sufficient to reduce SARS-CoV-2 replication in these cells. Further mechanistic studies are warranted to identify host factors[39,435] that contribute to varying disease severity during the course of COVID-19, along with the regulation of inflammatory and anti-inflammatory cellular processes in

SARS-CoV-2-infected cells.

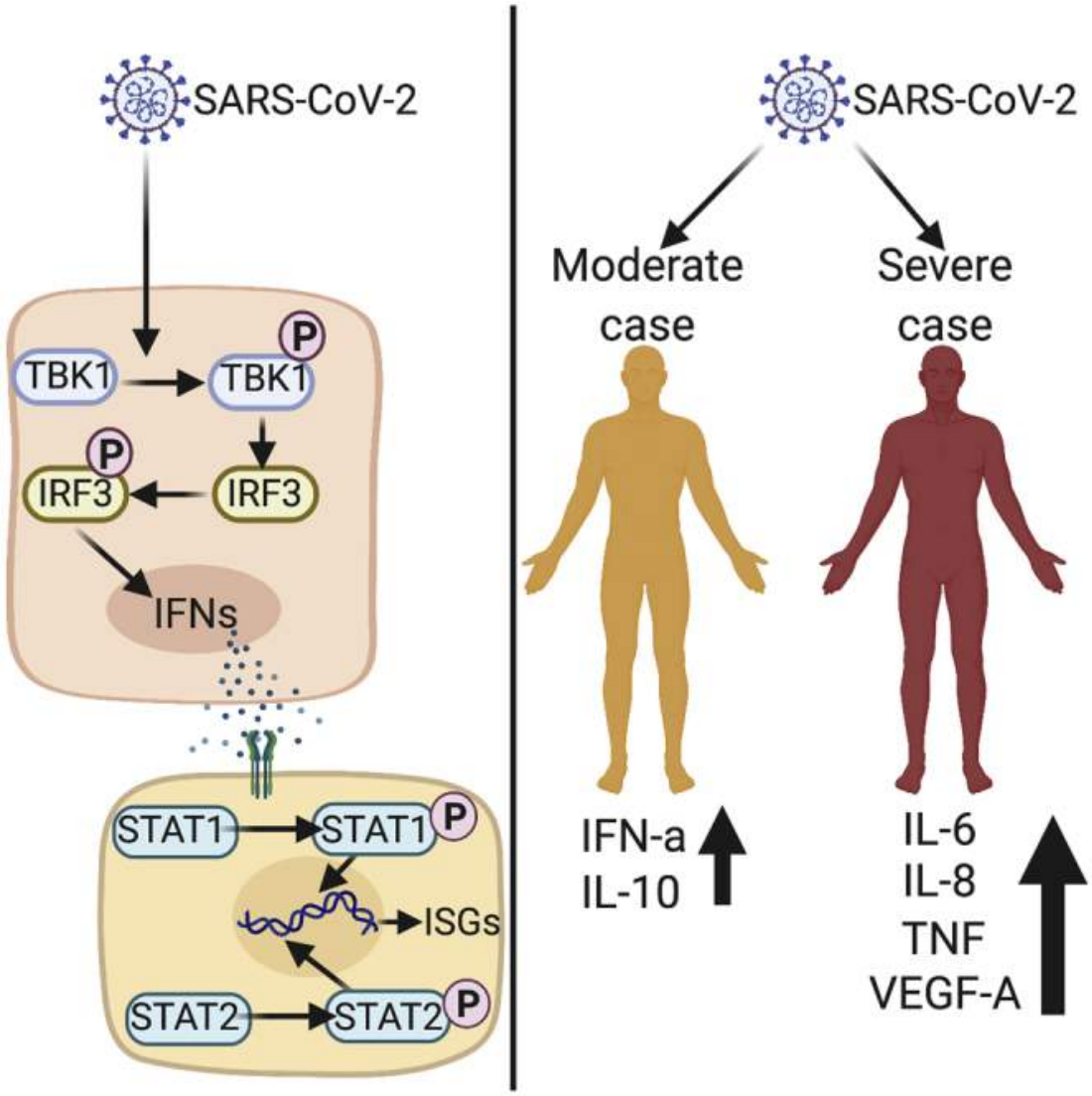


Figure 4.8: Depiction of proposed SARS-CoV-2-induced interferon response mechanism

4.2.2 Methods

Cells and viruses

Cell culturing and viral isolation performed by Dr. Arinjay Banerjee at McMaster University

Vero E6 cells (epithelial kidney cells isolated from the African green monkey; ATCC) were maintained in Dulbecco's modified Eagle's media (DMEM) supplemented with 10% fetal bovine serum (FBS; Sigma-Aldrich), 1x L-Glutamine and Penicillin/Streptomycin (Pen/Strep; VWR). Calu-3 cells (human lung adenocarcinoma derived; ATCC) were cultured as previously mentioned.[10] THF cells (human telomerase life-extended cells; from Dr. Victor DeFilippis' lab) were cultured as previously described.[37] Drosophila S2 cells (ThermoFisher Scientific) were cultured in Schneider's Drosophila medium supplemented with 10% FBS (Sigma-Aldrich) as recommended by the manufacturer and cells were incubated at 28°C. Stocks of genetically engineered vesicular stomatitis virus (VSV-GFP) carrying a green fluorescent protein (GFP) cassette (Noyce *et al.*, 2011) were stored at -80°C. H1N1 (A/Puerto Rico/8/1934 mNeon – 2A-HA) stocks were obtained from Dr. Matthew Miller's laboratory. HSV-GFP stocks were generated and maintained as mentioned previously.[261] Clinical isolate of SARS-CoV-2 (SARS-CoV-2/SB3) was propagated on Vero E6 cells and validated by next generation sequencing.[35] Virus stocks were thawed once and used for an experiment. A fresh vial was used for each experiment to avoid repeated freeze-thaws. VSV-GFP, HSV-GFP and H1N1 infections were performed at a multiplicity of infection (MOI) of 1. SARS-CoV-2 infections were performed at an MOI of 1 or 2. Experiments with SARS-CoV-2 were performed in a BSL3 laboratory and all procedures were approved by institutional biosafety committees at McMaster University and the University of Toronto.

Subject details

Acute patient sera (< 21 days from symptom onset) were acquired from moderate (hospital admission, but no ICU admission) and severe (ICU admission or death) cases of COVID-19 in Toronto, Canada, along with samples from uninfected, healthy individuals (see Table 15 for details). Work with patient sera was approved by the Sunnybrook Research Institute Research Ethics Board (amendment to 149-1994, March 2, 2020).[271]

RNA-Seq

RNA preparation performed by Dr. Arinjay Banerjee, Nader El-Sayes, Patrick Budylowski, and Rajesh Jacob at McMaster University

RNA was isolated from cells using RNeasy Mini kit (Qiagen). Sequencing was conducted at the McMaster Genomics Facility, Farncombe Institute at McMaster University. Sample quality was first assessed using a Bioanalyzer (Agilent), then enriched (NEBNext Poly(A) mRNA Magnetic Isolation Module; NEB). Library preparations were conducted (NEBNext Ultra II Directional RNA Library Prep Kit; NEB) and library fragment size distribution was verified (Agilent TapeSection D1000; Agilent). Libraries were quantified by qPCR, pooled in equimolar amounts, and qPCR and fragment size distribution verification were conducted again. Libraries were then sequenced on an Illumina HiSeq 1500 across 3 HiSeq Rapid v2 flow cells in 6 lanes (Illumina) using a paired-end, 2×50 bp configuration, with onboard cluster generation averaging 30.8M clusters per replicate (minimum 21.9M, maximum 46.0M).

Cytokine levels in COVID-19 patient sera

Sera preparation performed by Dr. Arinjay Banerjee, Nader El-Sayes, Patrick Budylowski, and Rajesh Jacob at McMaster University

Sera were analyzed using a 48-plex human cytokine and chemokine array by the manufacturer (Evetechnologies). Samples with an observed cytokine concentration (pg/mL) below the limit of detection (OOR <) were floored to the lowest observed concentration for that cytokine. Average $\log_2$FC for moderate patients ($n = 10$) versus healthy patients ($n = 5$) and severe patients ($n = 10$) versus healthy patients ($n = 5$) was plotted using the pheatmap R package (version 3.2.1) for all of the 48 cytokines. Cytokine expression levels were tested for significant differences via unpaired Student's t tests with Benjamini-Hochberg multiple testing correction using the stats R package (version 3.6.1).

Poly(I:C) transfection and IFN treatment

Cell transfection and interferon treatment performed by Dr. Arinjay Banerjee at McMaster University

Calu-3 cells were mock transfected with 4 µL or 8 µL of lipofectamine 3000 (ThermoFisher Scientific) in Opti-MEM (ThermoFisher Scientific) only or transfected with varying concentrations of poly(I:C) (InvivoGen) or poly(I:C)-rhodamine (InvivoGen). Recombinant human IFNβ1 was generated using Drosophila Schneider 2 (S2) cells following manufacturer's recommendation and by using ThermoFisher Scientific's Drosophila Expression system (ThermoFisher Scientific). Recombinant IFNβ1 was collected as part of the cell culture supernatant from S2 cells, and total protein was measured using Bradford assay.

Total protein concentration was used for subsequent experiments. To demonstrate that S2 cell culture media did not contain non-specific stimulators of mammalian antiviral responses, recombinant green fluorescent protein (GFP) was also generated using the same protocol and used the highest total protein concentration (2 mg/mL) for mock-treated cells (Figure 12 B). S2 cell culture supernatant containing GFP did not induce an antiviral response in human cells (Figure 12 B). For VSV-GFP, HSV-GFP, and H1N1-mNeon infections, cells were treated with increasing concentrations of IFNβ1 or GFP (control) containing cell culture supernatant. SARS-CoV-2-infected cells were treated with supernatant containing IFNβ1 or GFP. Commercially bought recombinant IFN-α2 (Sigma-Aldrich) was used for experiments that utilized IFN-α2.

Quantitative PCR

Quantitative PCR performed by Nader El-Sayes and Rajesh Jacob at McMaster University

Calu-3 cells were seeded at a density of 3 $\times$ 105 cells/well in 12-well plates. Cells were infected with SARS-CoV-2 for 12 hours. 12 hours post incubation, mock infected or infected cells were mock stimulated or stimulated with poly(I:C) or IFNβ for 6 hours. RNA extraction was performed using RNeasy® Mini Kit (Qiagen) according to manufacturer's protocol 6 hours post poly(I:C) tranfection or . 200 ng of purified RNA was reverse transcribed using iScript™ gDNA Clear cDNA Synthesis Kit (Bio-Rad). Quantitative PCR reactions were performed with TaqMan™ Universal PCR Master Mix (ThermoFisher Scientific) using pre-designed Taqman gene expression assays (ThermoFisher Scientific) for IFNβ1 (catalog no. #4331182), IRF7 (catalog no. #4331182), IFIT1 (catalog no. #4331182) and GAPDH (catalog no. #4331182) according to manufacturer's protocol. Relative mRNA expression was normalized to GAPDH and presented as 1/ΔCt. To quantify SARS-CoV-2 genome levels, primers were designed to amplify a region (UpE) between ORF3a and E genes. Primer sequences used were SARS2 UpE F – ATTGTTGATGAGCCTGAAG and SARS2 UpE R–TTCGTACTCATCAGCTTG. qPCR to determine UpE levels was performed using SsoFast EvaGreen supermix (Bio-Rad) as previously described.[36]

Agarose gel electrophoresis

Gel electrophoresis performed by Nader El-Sayes at McMaster University

UpE qPCR gene products were also run on agarose gels (Invitrogen) as previously mentioned to visualize qPCR amplicons.[37]

Immunoblots

Immunoblots performed by Dr. Arinjay Banerjee and Wael Demian at McMaster University

Calu-3 cells were seeded at a density of 3 × 105 cells/well in 12-well plates. Cells were infected with SARS-CoV-2 at an MOI of 1. Control cells were sham infected. 12 to 24 hours post incubation, cells were transfected or treated with poly(I:C) or IFNβ, respectively for indicated times. Cell lysates were harvested for immunoblots and analyzed on reducing gels as mentioned previously.[37] Briefly, samples were denatured in a reducing sample buffer and analyzed on a reducing gel. Proteins were blotted from the gel onto polyvinylidene difluoride (PVDF) membranes (Immobilon, EMD Millipore) and detected using primary and secondary antibodies. Primary antibodies used were: 1:1000 mouse anti-GAPDH (EMD Millipore; Catalogue number: AB2302; RRID: AB_10615768), 1:1000 mouse anti-SARS/SARS-CoV-2 N (ThermoFisher Scientific; Catalogue number: MA5-29981; RRID: AB_2785780, 1:1000 rabbit anti-IFIT1 (ThermoFisher Scientific; Catalogue number: PA3-848; RRID: AB_1958733), 1:1000 rabbit anti-β-actin (Abcam; Catalogue number: ab8227; RRID:), 1:1000 rabbit anti-IRF3 (Abcam; Catalogue number: ab68481; RRID: AB_11155653), 1:1000 rabbit anti-pIRF3-S386 (Cell Signaling; Catalogue number: 4947; RRID: AB_823547), 1:1000 rabbit anti-TBK1 (Abcam; Catalogue number: ab40676; RRID: AB_776632), 1:1000 rabbit anti-pTBK1-S172 (Abcam; Catalogue number: ab109272; RRID: AB_10862438), 1:1000 rabbit anti-STAT1 (Cell Signaling; Catalogue number: 9172; RRID: AB_2198300), 1:1000 rabbit anti-pSTAT1-Y701 (Cell Signaling; Catalogue number: 9167; RRID: AB_561284), 1:1000 rabbit anti-STAT2 (Cell Signaling; Catalogue number: 72604; RRID: AB_2799824), 1:1000 rabbit anti-pSTAT2-Y690 (Cell Signaling; Catalogue number: 88410S; RRID: AB_2800123). Secondary antibodies used were: 1:5000 donkey anti-rabbit 800 (LI-COR Biosciences; Catalogue number: 926-32213; RRID: 621848) and 1:5000 goat anti-mouse 680 (LI-COR Biosciences; Catalogue number: 925-68070; RRID: AB_2651128). Blots were observed and imaged using Image Studio (LI-COR Biosciences) on the Odyssey CLx imaging system (LI-COR Biosciences).

Immunofluorescent microscopy

Immunoflourescent microscopy performed by Patrick Budylowski at McMaster University

Calu-3 cells were infected with SARS-CoV-2 (MOI 1) for different times, followed by fixation in 10% neutral buffered formalin (Sigma) for 1 hour. After fixation, cells were washed, permeabilized, and stained as mentioned previously.[37] Primary antibodies used were mouse anti-SARS/SARS-CoV-2 N (ThermoFisher Scientific; Catalogue number: MA5-29981; RRID: AB_2785780) and human anti-SARS-CoV-2 N (GenScript; Catalogue number:

A02039S). Secondary antibodies used were goat anti-mouse Texas Red-X (ThermoFisher Scientific; Catalogue number: T-6390; RRID: AB_2556778) and rat anti-human FITC (BioLegend; Catalogue number: 410719; RRID: AB_2721575). Images were acquired using an EVOS M5000 imaging system (ThermoFisher Scientific).

Antiviral bioassay

Antiviral bioassay performed by Dr. Arinjay Banerjee at McMaster University

THF cells were pre-treated or mock treated with recombinant human IFNβ, followed by VSV-GFP, HSV-GFP or H1N1-mNeon infection at an MOI of 1. Infected cells were incubated at 37°C for 1 hour with gentle rocking every 15 minutes. After 1 hour, virus inoculum was aspirated and Minimum Essential Medium (MEM) with Earle's salts (Sigma) containing 2% FBS and 1% carboxymethyl cellulose (CMC; Sigma) was added on the cells. Cells were incubated for 19 hours at 37°C and green fluorescent protein (GFP) or mNeon levels were measured using a typhoon scanner (Amersham, Sigma).

Transcript quantification and differential expression analysis

Daniel Richard at Harvard University contributed to this work

Sequence read quality was checked with FastQC, with reads subsequently aligned to the human reference transcriptome (GRCh37.67) obtained from the ENSEMBL database,[183] indexed using the 'index' function of Salmon (version 0.14.0)[291] with a k-mer size of 31. Alignment was performed using the Salmon 'quant' function with the following parameters: "-l A –numBootstraps 100 –gcBias –validateMappings". All other parameters were left to defaults. Salmon quantification files were imported into R (version 3.6.1) using the tximport library (version 1.14.0)[356] with the 'type' option set to 'salmon'. Transcript counts were summarized at the gene-level using the corresponding transcriptome GTF file mappings obtained from ENSEMBL. Count data was subsequently loaded into DESeq2 (version 1.26.0)[243] using the 'DESeqdatasetFromTximport' function. In order to determine time/treatment dependent expression of genes, count data was normalized using the 'estimateSizeFactors' function using the default 'median ratio method' and output using the 'counts' function with the 'normalized' option.

For subsequent differential-expression analysis, a low-count filter was applied prior to normalization, wherein a gene must have had a count greater than 5 in at least three samples in order to be retained. Using all samples, this resulted in the removal of 12,980 genes for a final set of 15,760 used. Principal Component Analysis (PCA) of samples

across genes was performed using the 'vst' function in DESeq2 (default settings) and was subsequently plotted with the ggplot2 package in R.[412] Differential expression analyses were carried out with three designs: (A) the difference between infection/control status across all time points, (B) considering the effects of post-infection time (i.e., the interaction term between time and infection status) and (C) the difference between infection/control status at individual time points. (A) and (B) were performed using the 'DESeq' function of DESeq2 using all samples, with results subsequently summarized using the 'results' function with the 'α' parameter set to 0.05; p values were adjusted using the Benjamini-Hochberg FDR method,[43] with differentially expressed genes filtered for those falling below an adjusted p value of 0.05. For (C), infected/mock samples were subset to individual time points, with differential expression calculated using DESeq as described above. Additionally, given the smaller number of samples at individual time-points, differential-expression analysis was also performed with relaxation of the low-count filter described above, with results and p value adjustments performed as above.

Viral transcript quantification

Daniel Richard at Harvard University and Benjamin Tremblay at The University of Waterloo contributed to this work

Paired-end sequencing reads were mapped to CDS regions of the SARS-CoV-2 genomic sequence (Assembly ASM985889v3 - GCF_009858895.2) obtained from NCBI, indexed using the 'index' function of Salmon (version 0.14.0) (Patro *et al.*, 2017) with a k-mer size of 31. Subsequently, reads were aligned using the Salmon 'quant' function with the following parameters: "-l A –numBootstraps 100 –gcBias –validateMappings". All other parameters were left to defaults. Salmon quantification files were imported into R (version 3.6.1) using the tximport library (version 1.14.0)[356] with the 'type' option set to 'salmon'. All other parameters were set to default. Transcripts were mapped to their corresponding gene products via GTF files obtained from NCBI. Count data was subsequently loaded into DESeq2 (version 1.26.0) (Love *et al.*, 2014) using the 'DESeqdatasetFromTximport' function. Principal Component Analysis (PCA) of samples across viral genes was performed using the 'vst' function in DESeq2 (default settings) and was subsequently plotted with the ggplot2 package in R[412] (Figure 4.9 A). As viral transcript levels increased over time post-infection, non-normalized transcript counts were first converted to a $\log_2$ scale, and subsequently compared these across time-points (Figure 4.9 B and Table 16). To look at the changes in the expression of viral transcripts relative to total viral expression as a function of post-infection time, normalized transcript counts were used to perform differential-expression analysis with DESeq2. Results and p value adjustments were performed as described above.

In order to compare host/viral expression patterns, normalized transcript counts from infected samples were compared with either normalized or non-normalized viral transcript counts (from the same sample) across time-points. For each viral transcript ($n = 12$), all host genes ($n = 15{,}760$, after filtering described above) were tested for correlated expression changes across matched infected samples ($n = 18$, across 5 time-points) using Pearson's correlation coefficient (via the 'cor.test' function in R). Correlation test p values were adjusted across all-by-all comparisons using the Benjamini-Hochberg FDR method, and gene-transcript pairs at adjusted $p < 0.05$ were retained. To account for possible effects of cellular response to plate incubation, viral transcript abundance was averaged at each time-point and compared to host transcript abundance similarly averaged at each time-point for non-infected samples; correlation testing was done all-by-all for $n = 5$ data-points. Host genes that correlated with viral transcription in mock samples across time were removed from subsequent analyses; to increase stringency, mock correlation was defined using unadjusted $p < 0.05$. Host genes were sorted by correlation coefficient (with any given viral transcript), with the top 100 unique genes retained for visualization. Normalized host transcript counts were z-score transformed per-gene using the 'scale' function in R, with normalized/unnormalized viral transcript counts similarly transformed per-transcript. Resulting z-score expression heat maps were generated using the Complexheatmap library in R (version 2.2.0).[143] heat maps were generated for normalized/unnormalized viral transcript counts, given the different information revealed by absolute and relative viral expression patterns.

Viral genome mapping

Daniel Richard at Harvard University and Benjamin Tremblay at The University of Waterloo contributed to this work

Paired-end RNA-seq reads were filtered for quality control with Trim Galore! (version 0.6.4_dev)[220] and mapped to the SARS-CoV-2 reference sequence (NC_045512.2) with the Burrow-Wheeler Aligner,[236] using the BWA-MEM algorithm.[235] Output SAM files were sorted and compressed into BAM files using Samtools (version 1.10).[237] Read coverage visualization was performed from within the R statistical environment (version 4.0.0) (RCoreTeam, 2017) using the 'scanBam' function from the Rsamtools R package (version 1.32.0) to extract read coverage data and the ggplot2 R package (version 3.3.0)[412] to plot read coverage histograms (using 300 bins across the SARS-CoV-2 sequence).

Cellular pathway enrichment analysis

Hassaan Maan at the Vector Institute for Artificial Intelligence, Toronto contributed to this work

To determine cellular pathways that were associated with differentially expressed genes (DEGs), the ActivePathways R (version 1.0.1)[284] package was utilized to perform gene-set based pathway enrichment analysis. DEGs at each time point were treated as an independent set for enrichment analysis. Fisher's combined probability test was used to enrich pathways after p value adjustment using Holm-Bonferroni correction. Pathways of gene-set size less than 5 and greater than 1000 were excluded. Only pathways enriched at individual time-points were considered for downstream analysis; pathways enriched across combined time points as determined by ActivePathways Brown's p value merging method were filtered out.

Visualization of enriched pathways across time points was done using Cytoscape (version 3.8.0)[344] and the EnrichmentMap plugin (version 3.2.1),[259] as outlined by Reimand *et al.*[315] Up-to-date Gene-Matrix-Transposed (GMT) files containing information on pathways for the Gene Ontology (GO) Molecular Function (MF), GO Biological Process (BP)[92] and REACTOME[192] pathway databases were utilized with ActivePathways. Only pathways that were enriched at specific time points were considered. Bar plots displaying top ActivePathway GO terms and REACTOME enrichments for infection versus mock were plotted using the ggplot2 R package (version 3.2.1) for 1-, 2-, 3-, and 12-hour time points. Zero and 6-hour time points were omitted due to a lack of sufficient numbers of differentially expressed genes required for functional enrichment analysis.

Statistical Analysis

Statistical analyses for RNA-seq data were performed in R and are mentioned under the respective RNA-seq analyses sections. All other statistical calculations were performed in GraphPad Prism (version 8.4.2; www.graphpad.com) using two-tailed paired t-test. Significance values are indicated in the figures and figure legends. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$.

Data Availability

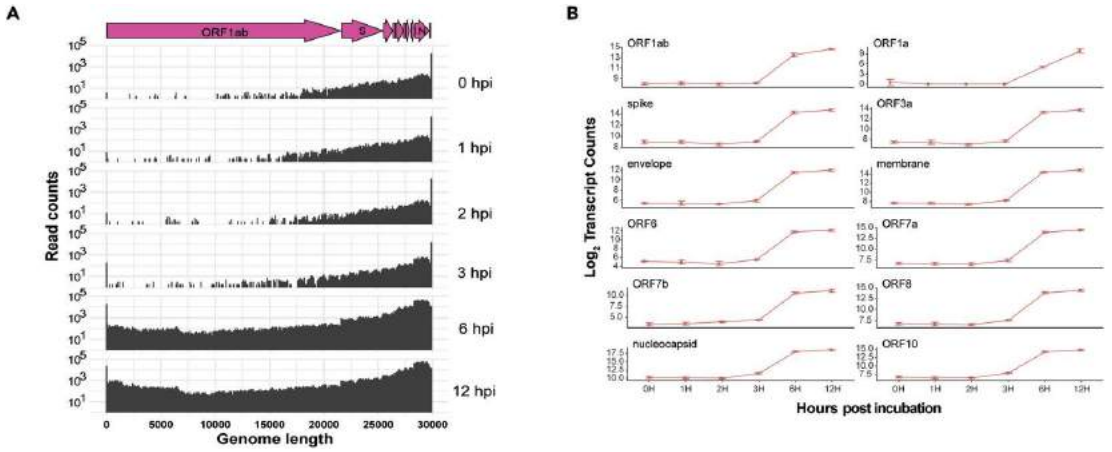
The DESeq2 normalized transcript counts for all genes with RNA-Seq data, significant or otherwise, plus the raw sequencing FASTQ reads have been deposited into the Gene Expression Omnibus (GEO) database with NCBI GEO accession number GSE151513.

4.2.3 Results

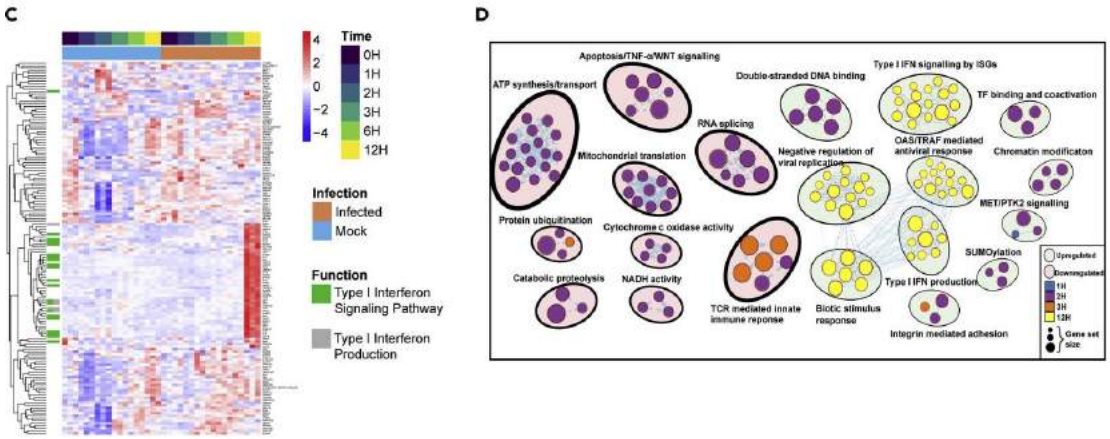
Global cellular response in SARS-CoV-2 infected human airway epithelial cells

The replication cycle of CoVs is complex and involves the generation of sub-genomic RNA molecules, which in turn code for mRNA that are translated into proteins.[33,333] To determine SARS-CoV-2 replication kinetics in human cells using RNA-seq, human airway epithelial cells (Calu-3) were infected at a multiplicity of infection (MOI) of 2. After incubation with virus inoculum for 1-hour, media was replaced with cell growth media and RNA was extracted and sequenced (poly-A enriched RNA) at 0-, 1-, 2-, 3-, 6- and 12-hours post incubation (hpi). SARS-CoV-2 genome, sub-genomic RNA and transcripts were detected in infected samples; viral transcript expression clustered based on post-incubation time using principal component analysis (PCA) (see Figure 10 A). In the RNA-seq analysis, high levels of expression of SARS-CoV-2 structural and accessory genes were detected at the 3' end of the genome as early as 0 hpi (Figure 4.9 A). Significant expression of *ORF1ab*, relative to 0 hpi was detected at 6 hpi (Figure 4.9 B). SARS-CoV-2 nucleocapsid (*N*) gene was highly expressed relative to other genes as early as 0 hpi (Figure 4.9 B), with relative expression significantly increasing over time ($p = 1.4e^{-16}$; Figure 4.9 B). The absolute expression of other genes increased over time with levels of $N > M > ORF10 > S > ORF1ab > ORF7a > ORF8 > ORF3a > ORF6 > E > ORF7b > ORF1a$ at 12 hpi (Figure 4.9 B and Table 16).

Figure 4.9

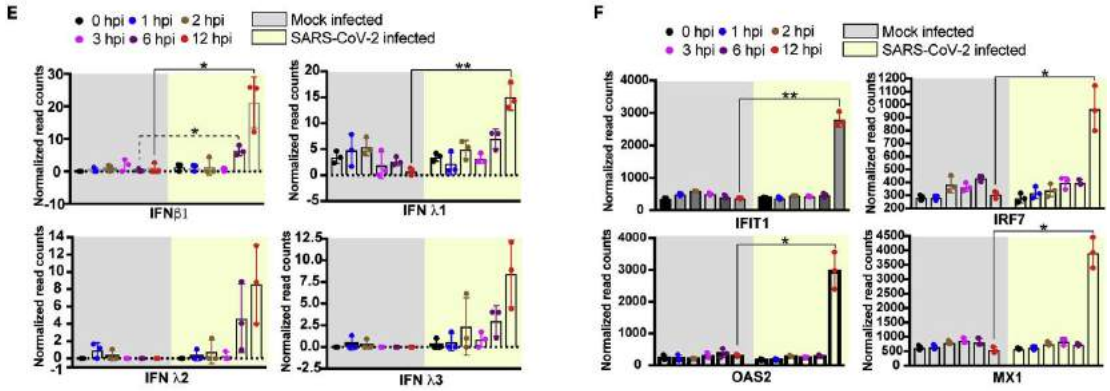


Calu-3 cells were infected with SARS-CoV-2 at an MOI of 1 or 2. RNA was extracted at different times post incubation. Viral and cellular gene expression was determined using time-series RNA-seq analysis or qPCR. (A) SARS-CoV-2 gene expression over 12 hours ($n = 3$/time point). The genome organization of SARS-CoV-2 is indicated above in pink.(B) Major SARS-CoV-2 gene expression levels at different times post incubation ($n = 3$/time point).

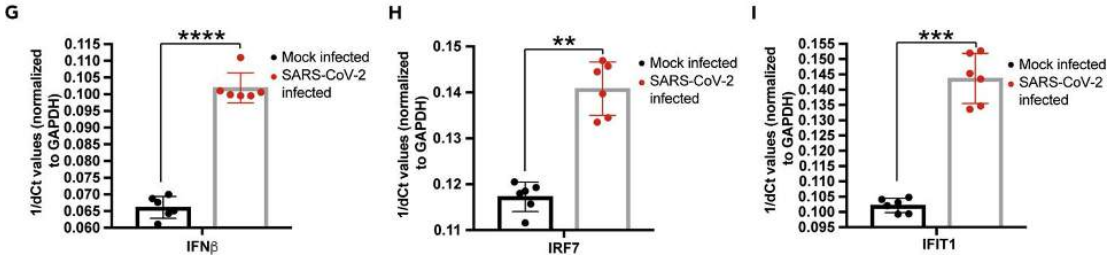


(C) Cellular genes ($n = 124$) that are significantly up- or down-regulated (FDR-adjusted $p < 0.05$; $|\log_2 FC| > 1$) in SARS-CoV-2-infected cells, relative to mock-infected cells at different times post incubation. Transcript levels are shown as z score normalized expression (scaled by gene). See Figure 10 E for a larger figure. (D) Cellular processes that are down- or up-regulated at different times post incubation. The size of the circles represents the number of genes that are down- or up-regulated at different times after incubation ($n = 3$/time point).

Figure 4.9



(E) Transcript abundance of type I and III interferon (IFN) genes (*IFNβ* and *IFNγ1-3*) in mock-infected and SARS-CoV-2-infected Calu-3 cells at different times ($n = 3$).(F) Transcript abundance of representative interferon-stimulated genes (ISGs) in mock-infected and SARS-CoV-2-infected Calu-3 cells at different times ($n = 3$).



(G) *IFNβ* transcript levels in Calu-3 cells infected with SARS-CoV-2 or mock infected for 12 hours, normalized to *GAPDH* ($n = 6$). Transcript levels were determined by qPCR.(H) *IRF7* transcript levels in Calu-3 cells infected with SARS-CoV-2 or mock infected for 12 hours, normalized to *GAPDH* ($n = 6$). Transcript levels were determined by qPCR. (I) *IFIT1* transcript levels in Calu-3 cells infected with SARS-CoV-2 or mock infected for 12 hours, normalized to *GAPDH* ($n = 6$). Transcript levels were determined by qPCR. Data are represented as mean ± SD, $n = 3$ or 6, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, and ****: $p < 0.0001$ (Student's t test). See also Figures 10–12, and Tables 14-16. Hpi, hours post incubation.

To determine SARS-CoV-2 infection-mediated host responses, total cellular RNA was extracted at different times post infection and analyzed gene expression in infected and mock infected Calu-3 cells using RNA-seq. Gene expression levels in these cells clustered based on time-points via PCA (see Figure 10 B). 124 genes were significantly differentially expressed in infected cells (FDR-adjusted $p < 0.05$), relative to mock infected cells in at

least one time point post infection ($|\log_2 FC| > 1$), including genes involved in type I IFN production and signaling (Figure 4.9 C; see Table 15 and Figure 10 C-E). The extent of antiviral gene expression at 12 hpi correlated with an increase in viral transcripts (see Figure 10 C). Interestingly, at early time points of 2 and 3 hpi, pathway enrichment analysis revealed numerous cellular processes that were significantly down-regulated in SARS-CoV-2 infected cells, relative to mock infected cells (FDR-adjusted $p < 0.05$). Down-regulated processes included RNA splicing, apoptosis, ATP synthesis and host translation, while genes associated with viral processes, cell adhesion and double-stranded RNA binding were up-regulated in infected cells relative to mock infected cells at 2 and 3 hpi (Figure 4.9 D; see Figures 10 D and 11, and Table 16). Cellular pathways associated with type I IFN production and signaling, along with OAS/TRAF-mediated antiviral responses were significantly up-regulated at 12 hpi (Figure 4.9 D and see Figure 11). Consistent with other reports,[52] transcript levels for *IFNβ1* and *IFNγ1* were significantly up-regulated at 12 hpi with SARS-CoV-2 (Figure 4.9 E). Transcript levels of *IFNγ2* and *IFNγ3* were elevated at 6 and 12 hpi, but the levels did not reach significance relative to mock infected cells at these time points (Figure 4.9 E).

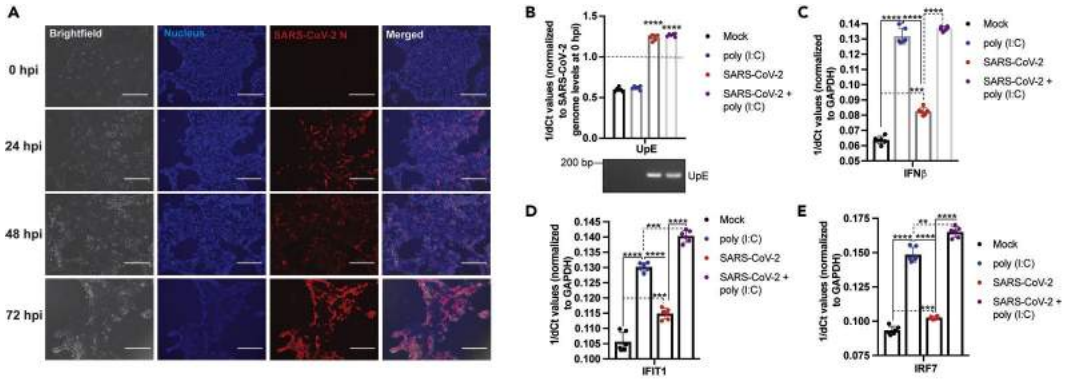
IFN production alone is not sufficient to protect cells from invading viruses. IFNs function through ISG expression, which in turn confers antiviral protection in infected (autocrine mode of action) and neighbouring (paracrine mode of action) cells.[337,338] Nineteen antiviral ISGs were up-regulated in infected cells, relative to mock infected cells at 12 hpi, including interferon induced protein with tetratricopeptide repeats 1 (*IFIT1*), interferon regulatory factor 7 (*IRF7*), 2'-5-oligoadenylate synthetase 2 (*OAS2*) and MX dynamin GTPase 1 (*MX1*) (Figure 4.9 F; see Figure 12 A and Table 13). Genes associated with structural molecule activity, cell adhesion and exocytosis were down-regulated in SARS-CoV-2 infected cells, relative to uninfected cells at 12 hpi (see Figure 11).

Coronaviruses, such as those that cause SARS and MERS have evolved multiple proteins that can inhibit type I IFN expression.[84,246,248,273,355,429] To confirm the RNA-seq findings that SARS-CoV-2 infection alone is sufficient to induce type I IFN and ISG responses in Calu-3 cells, cells were infected with SARS-CoV-2 and assessed transcript levels of *IFNβ*, *IRF7* and *IFIT1* by quantitative polymerase chain reaction (qPCR). *IFNβ* induction was observed 12 hpi in SARS-CoV-2 infected cells, relative to mock-infected cells (Figure 4.9 G). Consistent with the up-regulation of *IFNβ* transcripts in SARS-CoV-2 infected cells, transcript levels for ISGs, such as *IRF7* and *IFIT1* were also significantly up-regulated at 12 hpi relative to mock infected cells (Figure 4.9 H-I).

SARS-CoV-2 is not adept at inhibiting exogenous stimulation of type I IFN expression

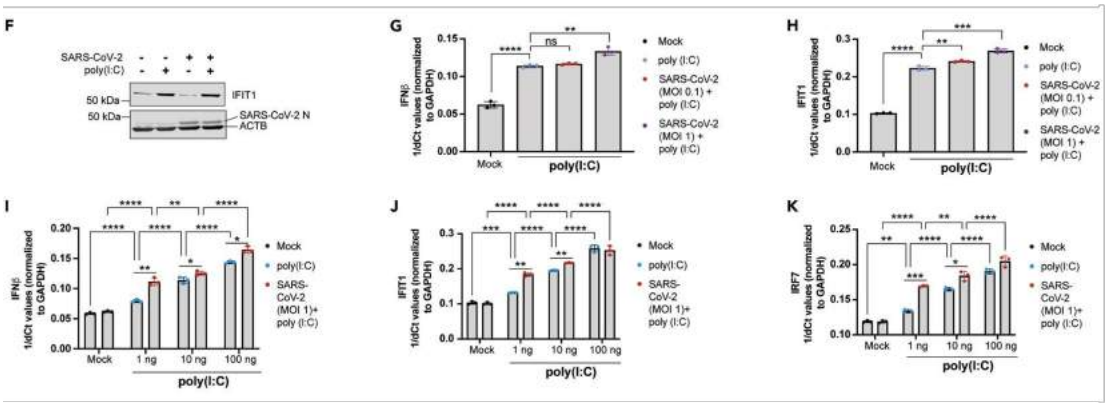
To determine if SARS-CoV-2 is able to inhibit type I IFN responses mounted against an exogenous stimulus, Calu-3 cells were infected with SARS-CoV-2 for 12 hours at an MOI of 1 and stimulated with exogenous double-stranded RNA [poly(I:C)] for 6 hours. SARS-CoV-2 replication was confirmed in Calu-3 cells over 0, 24, 48, and 72 hours of infection by staining for the nucleocapsid (N) protein (Figure 4.10 A). SARS-CoV-2 replication was quantified by qPCR using primers designed to amplify genomic RNA by targeting a region between *ORF3a* and *E* genes. This region was denoted as "upstream of E" (UpE). SARS-CoV-2 UpE levels were higher in SARS-CoV-2-infected cells and in SARS-CoV-2-infected + poly(I:C)-treated cells, relative to UpE levels at 0 hpi immediately after removing the inoculum (Figure 4.10 B). The levels of *IFNβ* transcripts in these cells were also measured by qPCR. Poly(I:C) transfection alone induced higher levels of *IFNβ* transcripts relative to mock-transfected cells (Figure 4.10 C). SARS-CoV-2 infection alone also induced higher levels of *IFNβ* transcripts relative to mock-infected cells (Figure 4.10 C). Interestingly, there was no significant difference in *IFNβ* transcript levels between poly(I:C)-transfected and SARS-CoV-2-infected + poly(I:C)-transfected cells (Figure 4.10 C).

Figure 4.10



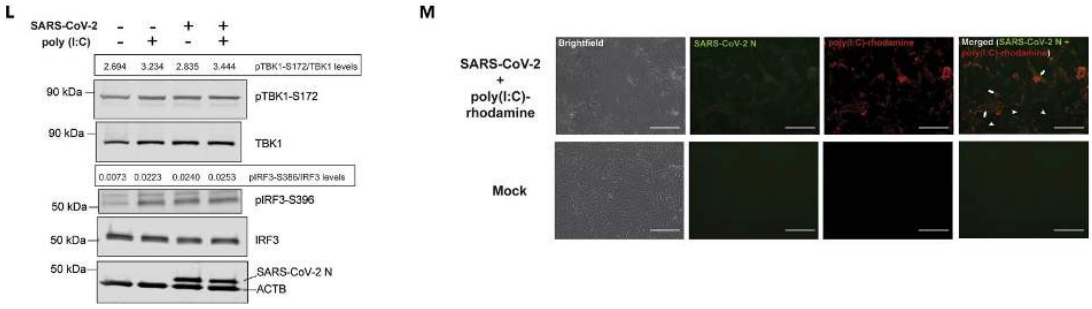
(A) Calu-3 cells were infected with SARS-CoV-2 (MOI 1) for 0, 24, 48, and 72 hours and stained to visualize the nucleus and SARS-CoV-2 nucleocapsid (N) protein. Scale bar indicates 300 µm.(B) SARS-CoV-2 genome (UpE) levels in Calu-3 cells infected with SARS-CoV-2 (MOI 1) or mock infected for 12 hours and transfected with 100 ng of poly(I:C) or mock transfected for 6 hours ($n = 6$). 1/dCT values are represented after normalizing Ct values for SARS-CoV-2 genome levels at 18 hpi with Ct values observed at 0 hpi. Gel (below): UpE qPCR amplicons were visualized on an agarose gel. (C) *IFNβ*, (D) *IFIT1*, or (E) *IRF7* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 12 hours. At 12 hpi, cells were either transfected with 100 ng of poly(I:C) or mock transfected for 6 hours. (C) *IFNβ*, (D) *IFIT1*, or (E) *IRF7* transcript levels were normalized to *GAPDH* transcript levels ($n = 6$).

Figure 4.10



(F) IFIT1, SARS-CoV-2 N, and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were either transfected with 1,000 ng of poly(I:C) or mock transfected for 24 hours ($n = 3$). (G) *IFNβ* or (H) *IFIT1* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 0.1 or 1) or mock infected for 24 hours. At 24 hpi, cells were transfected with 10 ng of poly(I:C) or mock transfected for 12 hours. (G) *IFNβ* or (H) *IFIT1* transcript levels were normalized to *GAPDH* transcript levels ($n = 3$). (I) *IFNβ*, (J) *IFIT1*, or (K) *IRF7* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were either transfected with varying concentrations of poly(I:C) or mock transfected for 12 hours. (I) *IFNβ*, (J) *IFIT1*, or (K) *IRF7* transcript levels were normalized to *GAPDH* transcript levels ($n = 3$).

Figure 4.10



(L) pTBK1-S172, TBK1, pIRF3-S396, IRF3, SARS-CoV-2 N, and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were either transfected with 1,000 ng of poly(I:C) or mock transfected for an additional 24 hours ($n = 3$). (M) Calu-3 cells were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours, followed by transfection with 1,000 ng of rhodamine-labeled poly(I:C) or mock transfection for 3 hours. Cells were fixed and stained to visualize SARS-CoV-2 nucleocapsid (N) protein and rhodamine-labeled poly(I:C). SARS-CoV-2 N and poly(I:C)-rhodamine containing cells are indicated by arrows. Cells that only contained detectable levels of poly(I:C)-rhodamine are indicated by arrow heads. Scale bar indicates 150 µm.Data are represented as mean ± SD, n = 3 or 6, **: $p < 0.01$, ***: $p < 0.001$, and ****: $p < 0.0001$ (Student's t test and Tukey's multiple comparisons test). pTBK1-S172 and pIRF3-S396 protein expression levels are expressed as ratios of pTBK1-S172/TBK1 and pIRF3-S396/IRF3 levels, respectively. Blots were quantified using Image Studio (Li-COR) ($n = 3$). Ct, cycle threshold. See also Figure 12.

To determine if *IFNβ* expression in SARS-CoV-2-infected and/or poly(I:C)-transfected cells is associated with ISG expression, the levels of *IFIT1* and *IRF7* were additionally quantified. Poly(I:C) transfection alone induced significantly higher levels of *IFIT1* and *IRF7* transcripts relative to mock-transfected cells (Figure 4.10 D-E). SARS-CoV-2 infection alone also induced higher levels of *IFIT1* and *IRF7* transcripts relative to mock-infected cells (Figure 4.10 D-E). Notably, *IFIT1* and *IRF7* transcript levels in SARS-CoV-2-infected + poly(I:C)-transfected cells were higher than levels in cells that were transfected with poly(I:C) alone (Figure 4.10 D-E), suggesting an additive effect of SARS-CoV-2 infection on poly(I:C)-mediated gene expression.

To validate gene expression observations, SARS-CoV-2 N, IFIT1, and beta-actin (ACTB) protein expression were examined. Poly(I:C) transfection induced higher levels of IFIT1 in Calu-3 cells, whereas SARS-CoV-2 infection did not induce higher observable levels of IFIT1 by immunoblot analysis at 48 hpi, relative to mock-infected cells (Figure 4.10 F); however, at 72 hpi, SARS-CoV-2 infection induced higher observable levels of IFIT1 protein

expression relative to mock-infected cells (see Figure 13). SARS-CoV-2 infection in these cells by detecting N protein in the samples (Figure 4.10 F).

To determine if the MOI of SARS-CoV-2 would influence its ability to modulate exogenous stimulation of interferon responses, Calu-3 cells were infected with two different MOIs of 0.1 and 1 for 24 hours, followed by exogenous stimulation of cells with 10 ng of poly(I:C) for 12 hours. Both MOIs of SARS-CoV-2 were unable to suppress the expression of *IFNβ* and *IFIT1* in poly(I:C)-stimulated cells (Figure 4.10 G-H). Furthermore, a high MOI of 1 had an additive effect on the expression levels of *IFNβ* and *IFIT1* in poly(I:C)-stimulated cells (Figure 4.10 G-H). Next, to determine if high concentrations of poly(I:C) in Figure 4.10 C–E may have overwhelmed the ability of SARS-CoV-2 to suppress IFN responses, Calu-3 cells were infected with SARS-CoV-2 for 24 hours, followed by stimulation with a range of concentrations of poly(I:C) for 12 hours (Figure 4.10 I–K). Even at the lowest poly(I:C) concentration of 1 ng, SARS-CoV-2 was unable to suppress *IFNβ*, *IFIT1*, and *IRF7* gene expression. Indeed, SARS-CoV-2 infection displayed an additive effect on the expression levels of *IFNβ* at all concentrations of poly(I:C), whereas the additive effect of SARS-CoV-2 infection on *IFIT1* and *IRF7* expression levels reached significance at concentrations of 1 ng and 10 ng of poly(I:C) (Figure 4.10 I–K).

Type I IFN production is primarily mediated by the phosphorylation and activation of TBK1, which in turn phosphorylates and activates IRF3.[189,204] Activation of TBK1 is associated with phosphorylation of serine 172,[226] whereas activation of IRF3 involves phosphorylation of serine 396, among other residues.[83] To determine SARS-CoV-2 infection-induced phosphorylation of TBK1 and IRF3, Calu-3 cells were infected for 24 hours followed by poly(I:C) or mock stimulation for another 24 hours and performed immunoblot analysis to detect levels of TBK1 (pTBK1-S172) and IRF3 (pIRF3-S396) phosphorylation. Only modest increases in phosphorylation of TBK1 were observed in SARS-CoV-2-infected and poly(I:C)-treated cells relative to untreated cells at the time of sampling (Figure 4.10 L). Phosphorylation of IRF3 was observed in both SARS-CoV-2-infected and poly(I:C)-treated cells relative to untreated cells, with similar levels of pIRF3-S396 observed following all infection and treatment conditions (Figure 4.10 L).

Titrating different concentrations of poly(I:C) in Figure 4.10 I–K demonstrated that SARS-CoV-2 infection has an additive effect on poly(I:C)-mediated up-regulation of IFN responses. In addition, it was also determined if poly(I:C) was delivered to infected cells (Figure 4.10 M). Calu-3 cells were infected with SARS-CoV-2 for 24 hours, followed by transfection with rhodamine-labeled poly(I:C) for 3 hours. At 24 hpi, visible levels of SARS-CoV-2 N and poly(I:C) could be detected in Calu-3 cells (Figure 4.10 M; arrows). Few uninfected cells (Figure 4.10 M; arrow heads) also contained detectable levels of poly(I:C); however, as identified in Figure 4.10 I–K, these cells are not sufficient to mount

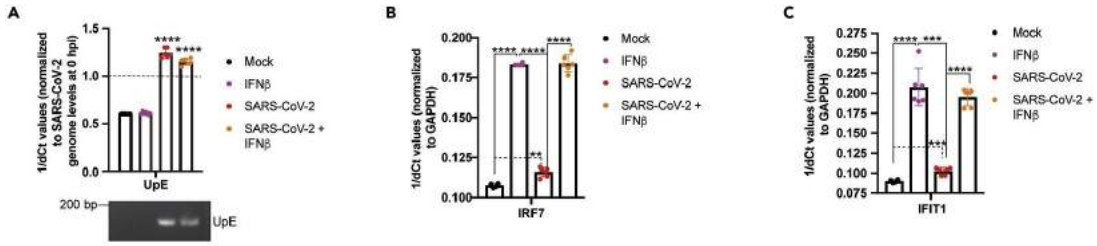
an overwhelming IFN response, because SARS-CoV-2 infection had an additive effect on IFN and ISG expression in poly(I:C)-treated cells.

SARS-CoV-2 infection is unable to suppress downstream type I IFN signaling

SARS-CoV and MERS-CoV proteins can also inhibit downstream IFN signaling to restrict the production of ISGs.[101] To evaluate if SARS-CoV-2 can inhibit type I IFN signaling in response to exogenous IFNβ treatment, Calu-3 cells were infected for 12 hours at an MOI of 1 and stimulated these cells with recombinant human IFNβ for 6 hours. Gene expression levels of *IRF7* and *IFIT1* were monitored in these cells by qPCR. Validation of the antiviral efficacy of the recombinant IFNβ1 was carried out in human fibroblast (THF) cells that were pretreated with IFNβ1, followed by RNA and DNA virus infections. Pre-treatment of THF cells with recombinant IFNβ1 inhibited the replication of herpes simplex virus (HSV), vesicular stomatitis virus (VSV), and H1N1 in a dose-dependent manner (see Figure 12 B).

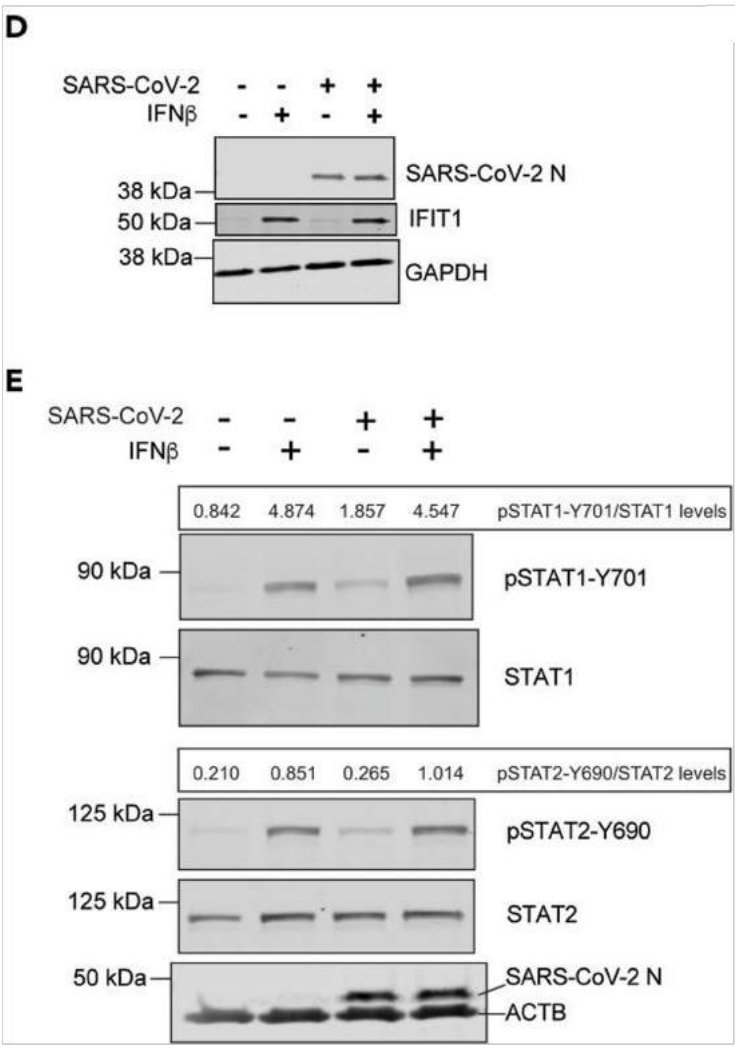
SARS-CoV-2 genome levels were significantly higher in infected cells relative to mock-infected cells (Figure 4.11 A). Although SARS-CoV-2 UpE levels displayed a lower trend in SARS-CoV-2-infected + IFNβ-treated cells relative to SARS-CoV-2-infected-only cells, UpE levels were not significantly different after 6 hours of IFNβ treatment (Figure 4.11 A). Exogenous IFNβ treatment significantly up-regulated transcript levels of *IRF7* and *IFIT1* relative to mock-treated Calu-3 cells (Figure 4.11 B-C). Consistent with this study's RNA-seq data, SARS-CoV-2 infection induced mild but significant levels of *IRF7* and *IFIT1* transcripts relative to mock-infected cells (Figure 4.11 B-C). IFNβ-mediated induction of *IRF7* and *IFIT1* was not dampened by SARS-CoV-2 infection (Figure 4.11 B-C).

Figure 4.11



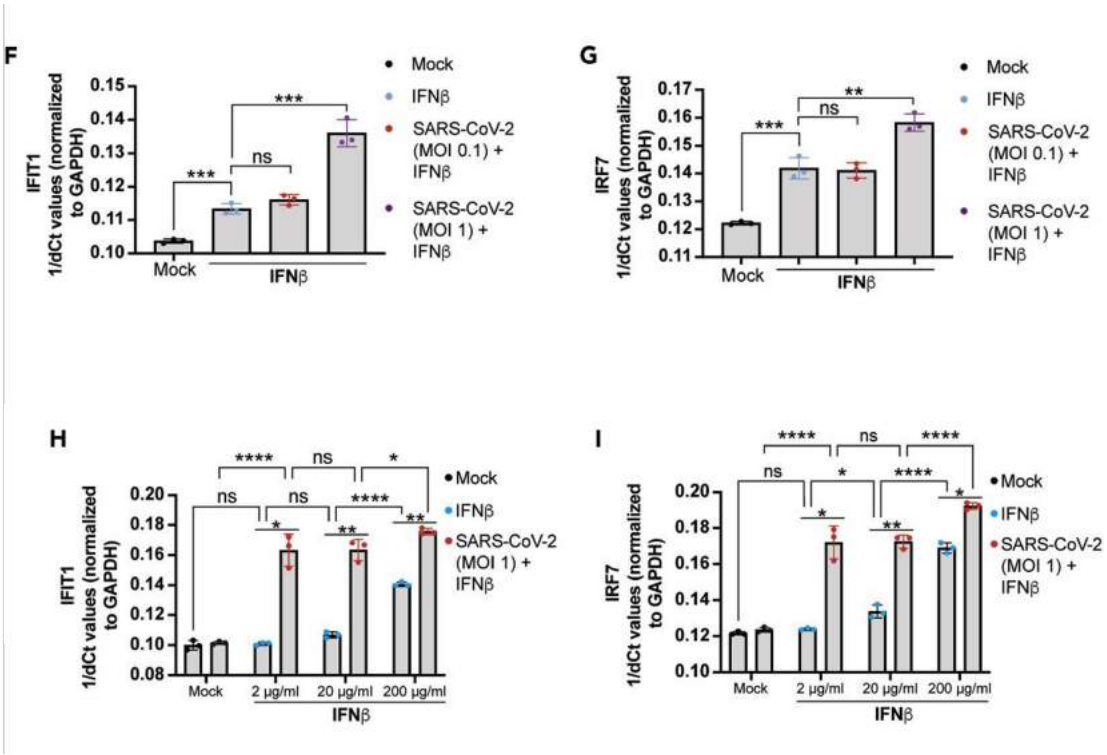
(A) SARS-CoV-2 genome (UpE) levels in Calu-3 cells infected with SARS-CoV-2 (MOI 1) or mock infected for 12 hours and treated with recombinant IFNβ or mock treated for 6 hours ($n = 6$). 1/dCT values are represented after normalizing Ct values for SARS-CoV-2 genome levels at 18 hpi with Ct values observed at 0 hpi. Gel (below): UpE qPCR amplicons were visualized on an agarose gel. (B) *IRF7* or (C) *IFIT1* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 12 hours. At 12 hpi, cells were either treated with recombinant IFNβ or mock treated for 6 hours. (B) *IRF7* or (C) *IFIT1* transcript levels were normalized to *GAPDH* transcript levels ($n = 6$).

Figure 4.11



(D) SARS-CoV-2 N, IFIT1, and GAPDH protein expression in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 12 hours. At 12 hpi, cells were either treated with recombinant IFNβ or mock treated for 6 hours ($n = 3$). (E) pSTAT1-Y701, STAT1, pSTAT2-Y690, STAT2, SARS-CoV-2 N, and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were either treated with recombinant IFNβ or mock treated for 30 minutes ($n = 3$).

Figure 4.11



(F) *IFIT1* or (G) *IRF7* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 0.1 or 1) or mock infected for 24 hours. At 24 hpi, cells were mock treated or treated with recombinant IFNβ containing media (20 µg/mL total protein) for 12 hours. (F) *IFIT1* or (G) *IRF7* transcript levels were normalized to *GAPDH* transcript levels ($n = 3$). (H) *IFIT1* or (I) *IRF7* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were either treated with varying concentrations of recombinant IFNβ or mock treated for 12 hours. (H) *IFIT1* or (I) *IRF7* transcript levels were normalized to *GAPDH* transcript levels ($n = 3$). Data are represented as mean ± SD, $n = 3$ or 6, ns: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, and ****: $p < 0.0001$ (Student's t test and Tukey's multiple comparison's test). Ct, cycle threshold. pSTAT1-Y701 and pSTAT2-Y690 protein expression levels are expressed as ratios of pSTAT1-Y701/STAT1 and pSTAT2-Y690/STAT2 levels, respectively. Blots were quantified using Image Studio (Li-COR) ($n = 3$). For IFNβ treatment, cell culture supernatant containing recombinant IFNβ was used. Cell culture supernatant containing 2 mg/mL of total protein, including IFNβ, was used in A-E. A range of concentrations was used for other figures as indicated. Ct, cycle threshold. See methods for recombinant IFNβ generation. See also Figures 12 and 13.

To validate the transcriptional responses, experiments with exogenous IFNβ treatment were repeated to determine if SARS-CoV-2 could inhibit type I IFN-mediated up-regulation of IFIT1 at the protein level. SARS-CoV-2 infection alone failed to induce detectable levels of IFIT1 at 12 hpi (Figure 4.11 D). IFNβ treatment with or without prior 12 hours of SARS-CoV-2 infection induced robust expression of IFIT1 (Figure 4.11 D). SARS-CoV-2 infection was confirmed in these cells by immunoblotting for N protein (Figure 4.11 D).

Binding of IFNs to their receptors activates a series of downstream signaling events, which involves phosphorylation of STAT1 at tyrosine 701 (pSTAT1-Y701) and STAT2 at tyrosine 690 (pSTAT2-Y690).[302,363] To determine if SARS-CoV-2 can inhibit phosphorylation of STAT1 and STAT2 proteins, Calu-3 cells were infected with SARS-CoV-2 for 24 hours followed by 30 minutes of stimulation with or without recombinant IFNβ. SARS-CoV-2 infection alone induced mild pSTAT1-Y701 and pSTAT2-Y690 levels relative to mock-infected cells, albeit lower than levels observed in exogenous IFNβ-treated cells (Figure 4.11 E). Importantly, SARS-CoV-2 infection was unable to inhibit pSTAT1-Y701 and pSTAT2-Y690 levels in cells treated with IFNβ (Figure 4.11 E).

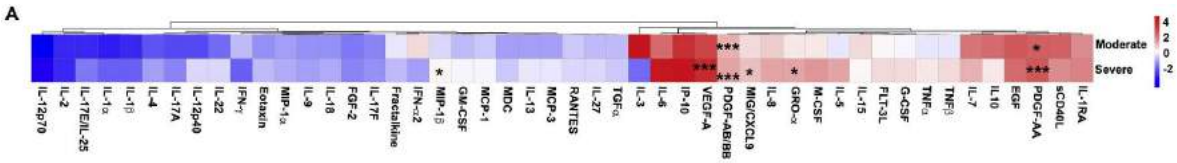
To determine if the MOI of SARS-CoV-2 would influence its ability to suppress exogenous stimulation of interferon responses, Calu-3 cells were infected with two different MOIs of 0.1 and 1 for 24 hours, followed by exogenous stimulation of cells with IFNβ for 12 hours. Infection with both MOIs of SARS-CoV-2 was unable to suppress the expression of *IFIT1* and *IRF7* on IFNβ treatment (Figures 4.11 F-G). Furthermore, high MOI 1 of SARS-CoV-2 had an additive effect on the expression of *IFIT1* and *IRF7* in IFNβ-treated cells (Figures 4.11 F-G). Next, to determine if high concentrations of IFNβ in Figures 4.11 A-D may have overwhelmed the ability of SARS-CoV-2 to suppress IFN-mediated responses, Calu-3 cells were infected with SARS-CoV-2 for 24 hours, followed by stimulation with a range of concentrations of IFNβ for 12 hours (Figures 4.11 H-I). SARS-CoV-2 was unable to suppress *IFIT1* and *IRF7* gene expression. MOI 1 of SARS-CoV-2 had an additive effect on the expression of *IFIT1* and *IRF7* in IFNβ-treated cells at all concentrations of IFNβ (Figures 4.11 H-I).

Cytokine levels in COVID-19 patients and effect of type I IFNs on SARS-CoV-2 replication

To evaluate type I IFN and other infection-associated cytokines in COVID-19 patients, acute sera ($<$ 21 days from symptom onset) was analyzed from 20 COVID-19 positive patients, of whom 10 were categorized as "moderate" cases requiring hospital admission, but not admission to intensive care unit (ICU). The remaining 10 samples were from "severe" cases that required ICU admission. For severe cases, 6/10 patients died, and 10/10 moderate

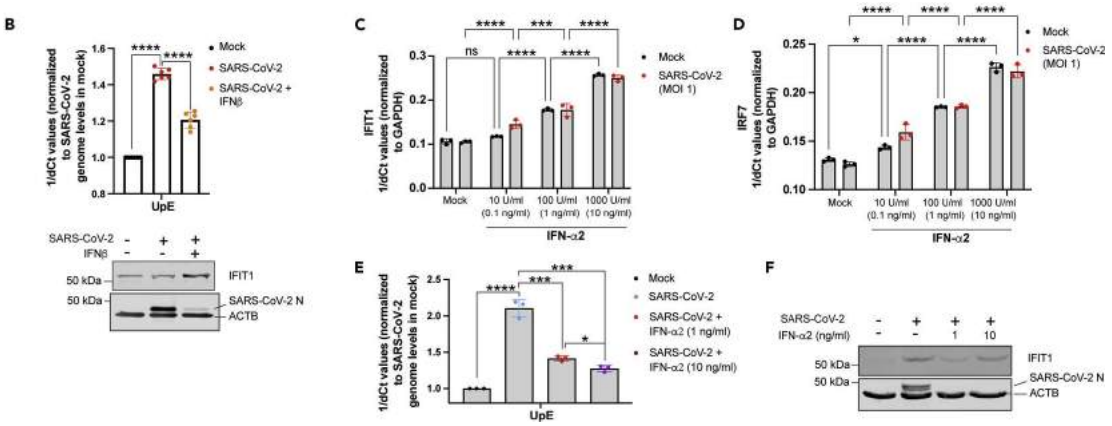
cases were discharged (see Table 15). Sera from 5 healthy, uninfected individuals was also included. Sera from moderate cases of COVID-19 displayed significantly higher levels of platelet-derived growth factor AA (PDGF-AA) and PDGF-AB/BB relative to uninfected individuals (Figure 4.12). Patients with severe COVID-19 displayed significantly higher levels of PDGF-AA, PDGF-AB/BB, GROα (CXCL-1), CXCL-9, MIP-1β and vascular endothelial growth factor A (VEGF-A) relative to healthy individuals (Figure 4.12). In addition, severe cases of COVID-19 displayed an increasing trend for levels of interleukin-6 (IL-6), IL-5, macrophage colony stimulating factor 1 (M-CSF), IL-8, tumor necrosis factor alpha (TNFα), TNFβ, and granulocyte colony stimulating factor 1 (G-CSF) relative to healthy individuals and moderate cases of COVID-19. In addition, both moderate and severe cases of COVID-19 displayed an increasing trend for IL-7 and IP-10 relative to healthy controls, although the data were not significant due to wide within-patient variation in acute serum samples. Moderate cases of COVID-19 displayed an increasing trend for levels of IFN-α2 and IL-10 relative to healthy individuals and severe cases of COVID-19 (Figure 4.12 and see Tables 15 and 16).

Figure 4.12



A) Protein levels were analyzed in sera using a 48-plex human cytokine and chemokine array. Mean $\log_2$ fold-change in serum cytokine protein levels in patients with moderate ($n = 10$) or severe ($n = 10$) case of COVID-19, relative to levels in healthy donors ($n = 5$) are represented here.

Figure 4.12



(B) SARS-CoV-2 genome (UpE) levels in Calu-3 cells infected with SARS-CoV-2 (MOI 1) or mock infected for 1 hour followed by treatment with recombinant IFNβ or mock treatment for 72 hours ($n = 6$). 1/dCT values are represented after normalizing Ct values for SARS-CoV-2 genome levels in infected cells (with or without recombinant IFNβ treatment) with Ct values observed in mock-infected cells. Blot (below): IFIT1, SARS-CoV-2 N, and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 or mock infected for 1 hour, followed by treatment with recombinant IFNβ or mock treatment for 72 hours ($n = 3$). (C) *IFIT1* or (D) *IRF7* transcript levels in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 24 hours. At 24 hpi, cells were treated with varying concentrations of recombinant IFN-α2 or mock treated for 6 hours. (C) *IFIT1* or (D) *IRF7* transcript levels were normalized to *GAPDH* transcript levels ($n = 3$). (E) SARS-CoV-2 genome (UpE) levels in Calu-3 cells infected with SARS-CoV-2 (MOI 1) or mock infected for 1 hour followed by treatment with recombinant IFN-α2 (1 ng/mL or 10 ng/mL) or mock treatment for 72 hours ($n = 3$). 1/dCT values are represented after normalizing Ct values for SARS-CoV-2 genome levels in infected cells (with or without recombinant IFN-α2 treatment) with Ct values observed in mock-infected cells. (F) IFIT1, SARS-CoV-2 N, and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 (MOI 1) or mock infected for 1 hour, followed by treatment with recombinant IFN-α2 (1 ng/mL or 10 ng/mL) or mock treatment for 72 hours ($n = 3$). Data are represented as mean $\pm$ SD, $n = 5$ for healthy human controls, and $n = 10$ each for moderate or severe cases of COVID-19, $n = 3$ or 6 for *in vitro* experiments. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, and ****: $p < 0.0001$ (Student's t tests with Benjamini-Hochberg multiple testing correction, Student's t tests and Tukey's multiple comparison test). See also Tables 15 and 16, and Figure 12.

To determine if exogenous IFNβ treatment can inhibit SARS-CoV-2 replication, Calu-3 cells were infected for 1 hour, following which they were either mock treated or treated with recombinant IFNβ for 72 hours. Exogenous IFNβ treatment reduced SARS-CoV-2 genome

(UpE) and N protein levels in these cells (Figure 4.12 B), consistent with an increase in IFIT1 levels (Figures 4.12 B and 13).

Next, to determine if levels of IFN-α2 that were detected in sera from patients who developed moderate COVID-19 (Figure 4.12 A and see Table 16) were sufficient to induce an IFN response, a range of concentrations of IFN-α2 were tested against SARS-CoV-2. Calu-3 cells contained higher levels of *IFIT1* transcripts in the presence of medium (1 ng/mL) and high (10 ng/mL) concentrations of IFN-α2, whereas *IRF7* transcript levels were higher in Calu-3 cells treated with low (0.1 ng/mL), medium, or high concentrations of IFN-α2 (Figure 4.12 C-D). Furthermore, consistent with the data for IFNβ, SARS-CoV-2 infection was unable to suppress IFN-α2-mediated expression of ISGs, such as *IFIT1* and *IRF7* (Figure 4.12 C-D). Finally, to determine if IFN-α2 was capable of suppressing SARS-CoV-2 replication, Calu-3 cells were infected with SARS-CoV-2 and treated the cells with two concentrations of IFN-α2 (1 ng/mL and 10 ng/mL) for 72 hours. Both concentrations of IFN-α2 significantly reduced SARS-CoV-2 replication (Figure 4.12 E-F).

4.2.4 Discussion

SARS-CoV-2 emerged in December 2019 to cause a pandemic of COVID-19.[105,440] Clinical observations and emerging data from *in vitro* and *in vivo* studies have demonstrated the ability of SARS-CoV-2 to induce type I IFNs.[52,314,430] However, a recent review summarized studies that suggest that antiviral IFN responses are dampened in COVID-19 patients.[8] Emerging data also suggest that timing and extent of interferon production is likely associated with manifestation of disease severity.[441] In spite of some progress in understanding how SARS-CoV-2 activates antiviral responses, mechanistic studies into SARS-CoV-2-infection-mediated induction and modulation of human type I IFN responses are lacking. To understand SARS-CoV-2-infection-induced pathogenesis during the clinical course of COVID-19, it is imperative that we understand if and how replicating SARS-CoV-2 interacts with type I IFN responses. These observations can be leveraged to develop drug candidates and inform ongoing drug trials, including trials that involve type I and III IFNs.

In this study, a time-series RNA-seq analysis of poly(A))-enriched RNA from SARS-CoV-2-infected human airway epithelial cells allowed us to map the progression of SARS-CoV-2 replication and transcription. As observed with other coronaviruses,[114,225,299] SARS-CoV-2 replicated and transcribed sub-genomic RNA and mRNA in a directional manner (Figure 4.9 A-B). Thus, this study's data demonstrate that SARS-CoV-2 replication strategy is consistent with other coronaviruses. Furthermore, the data in this study demonstrate that Calu-3 cells support SARS-CoV-2 replication and that these cells represent a good *in vitro* model to study SARS-CoV-2-host interactions, as reported by others.[52,430]

Studies have demonstrated that ectopically expressed SARS-CoV-2 proteins can suppress type I IFN responses. Low SARS-CoV-2-induced type I IFN responses may be associated with (1) the virus' ability to mask the detection of viral RNA by cellular PRRs and/or (2) the ability of viral proteins to inactivate cellular mechanisms involved in type I IFN induction.[347] Data from this study show that infection with wild-type and replication competent SARS-CoV-2 is capable of inducing a type I IFN response in human airway epithelial cells, including up-regulation of IFN expression (Figure 4.9 C, E, G) and downstream IFN signaling genes (Figures 4.9 C, F, H, I and 11). These observations corroborate and expand upon recent data from Lei, Rebendenne, and Yin *et al.*'s studies where the authors have demonstrated that SARS-CoV-2 infection is capable of up-regulating type I IFN responses in multiple human cell types.[230,314,430]

The physiological relevance of an existing but dampened type I IFN response to SARS-CoV-2 remains to be identified. Emerging data suggest that prolonged and high levels of type I IFNs correlate with COVID-19 disease severity.[247] Thus, a dampened yet protective early type I IFN response against SARS-CoV-2 may in fact be beneficial for humans.[287] However, questions remain about how a low-type I IFN response against SARS-CoV-2 could play a protective role during infection. One possibility is that low levels of type I IFN production are sufficient to control SARS-CoV-2 replication (Figure 4.12 E). This may explain the large number of asymptomatic cases of SARS-CoV-2 where an early IFN response may control virus replication and disease progression. Indeed, in one study, type I IFN (IFNα) levels were higher in asymptomatic cases relative to symptomatic cases ($n = 37$).[241] Further studies are required to identify regulatory mechanisms behind the protective role of a controlled and early IFN response during SARS-CoV-2 infection versus the delayed and potentially damaging long-term IFN response observed in some severe cases of COVID-19.

In one study, SARS-CoV was demonstrated to inhibit poly(I:C)-mediated up-regulation of IFNβ.[246] The data in this study show that infection with SARS-CoV-2 is unable to inhibit poly(I:C)-mediated up-regulation of *IFNβ* transcripts and downstream ISGs, such as *IFIT1* and *IRF7* (Figure 4.10 C-E). Indeed, SARS-CoV-2 infection, followed by poly(I:C) transfection induced higher levels of *IFNβ* and ISG (*IFIT1* and *IRF7*) transcripts relative to poly(I:C) alone, indicating that wild-type infection partially augments poly(I:C)-mediated up-regulation of type I IFN signaling (Figure 4.10 C-E, G-K) in Calu-3 cells. Furthermore, low or high MOI of SARS-CoV-2 was unable to suppress IFN responses stimulated by poly(I:C) (Figure 4.10 G-H). A high MOI of SARS-CoV-2 was also unable to suppress IFN responses stimulated by a range of poly(I:C) concentrations (Figure 4.10 I-K). Thus, it is important to identify the kinetics and landscape of virus infection, transcription and translation, and how that may regulate human type I IFN responses. Although multiple

studies have demonstrated that ectopically expressed SARS-CoV-2 proteins can suppress type I IFN responses, there is a need to study the dynamic interplay between viral RNA-mediated induction of IFN responses, followed by subsequent dampening of these responses, as viral proteins accumulate in infected cells. It is also important to identify SARS-CoV-2-host interactions in different human cell types to discern cell-type-specific differences in IFN responses. Furthermore, as multiple SARS-CoV-2 variants continue to evolve, it is important to assess the ability of these emerging variants to modulate human type I IFN responses.

Coronaviruses, including highly pathogenic SARS-CoV, MERS-CoV, and porcine epidemic diarrhea virus (PEDV) have evolved proteins that can efficiently inhibit type I IFN responses.[84,103,246,248,273,355,423,429] In spite of observing statistically significant up-regulation of type I IFNs and ISGs at 12 hpi with SARS-CoV-2 (Figure 4.9), in preliminary studies collaborators and I were unable to observe detectable levels of pIRF3-S396 prior to accumulation of antiviral mRNAs. My collaborators have previously shown that antiviral responses can be induced in the absence of prototypic markers of IRF3 activation such as dimerization and hyperphosphorylation, even when IRF3 was shown to be essential.[275] The simplest interpretation is that early activation of IRF3-mediated IFN responses requires low (or even undetectable) levels of pIRF3-S396, which accumulate to detectable levels over time (Figure 4.10 L).

SARS-CoV and MERS-CoV can inhibit phosphorylation and activation of STAT1 and STAT2, which blocks global IFN-induced antiviral responses.[101] The data in this study demonstrate that SARS-CoV-2 infection induces phosphorylation of STAT1 and STAT2 (Figure 4.11 E), along with up-regulation of ISGs, such as *IRF7* and *IFIT1* (Figure 4.11 B-C). In addition, SARS-CoV-2 infection is unable to inhibit the activation of STAT1 and STAT2 by exogenous type I IFN (Figure 4.11 E), along with the expression of downstream ISGs, such as *IRF7* and *IFIT1* (Figure 4.11 B-C and see Figure 13). Although SARS-CoV-2 infection alone induced low levels of type I IFN (Figures 4.9 E and 4.10 B), it was sufficient to activate STAT proteins (Figure 4.11 E) and downstream ISG expression (Figures 4.10 B-C, 4.11 B-C; see Figures 10 and 12). Thus, the dampened ability of SARS-CoV-2 to inhibit downstream type I IFN responses compared with other zoonotic CoVs extends support to my hypothesis that the pathogenic consequences of a dampened type I IFN response may be largely negated by the sensitivity of SARS-CoV-2 to this response. Indeed, in these studies, exogenous type I IFN (IFNβ1 and IFN-α2) treatment significantly reduced SARS-CoV-2 replication in human airway epithelial cells (Figure 4.12 B, E, F), consistent with a recent study that compared the susceptibility of SARS-CoV and SARS-CoV-2 to type I IFNs.[240] Recent studies have identified the role of an impaired type I IFN response in COVID-19 disease severity,[39,435] which support the conclusion that SARS-CoV-2 is capable

of inducing a type I IFN response, and perhaps the inability of the host to mount this response contributes to disease severity. In addition, this study's data provide promising support for ongoing clinical trials that include type I IFN treatment.

Studies have demonstrated that COVID-19 patients mount a dysregulated immune response, which is associated with a poor clinical outcome.[247] In this study, it is observed that patients with moderate or severe case of COVID-19 had elevated serum levels of growth factors PDGF-AA and PDGF-AB/BB relative to healthy controls (Figure 4.12 A and Tables 15 and 16). The role of PDGFs in driving disease pathology has been described previously,[22] and therapeutic use of PDGF antagonists has also been recommended.[141,330] PDGF-BB has also been introduced in the clinic as a wound-healing therapy.[425] The physiological impact of elevated PDGF levels and cellular factors that regulate the expression of PDGF in COVID-19 patients remains to be understood.

Sera from patients with moderate case of COVID-19 contained higher levels of IL-10 relative to severe cases, which is suggestive of an anti-inflammatory response[94,308] (Figure 4.12 A and Tables 15 and 16). On the contrary, sera from patients with severe case of COVID-19 displayed a higher trend for levels of IL-6, IL-8, and TNFα relative to moderate cases, which is suggestive of a pro-inflammatory response (Figure 4.12 A and Tables 15 and 16).[247,251,308] Observations from this study (Figure 4.12 A), along with other recent reports,[241,247] warrant further investigations into mechanistic regulation of pro- and anti-inflammatory processes in SARS-CoV-2-infected human airway cells. Identifying regulatory proteins, such as transcription factors that contribute to a pro-inflammatory cytokine response or "cytokine storm" in SARS-CoV-2-infected individuals, will inform the selection and utilization of anti-inflammatory drugs.

Patients with moderate COVID-19 demonstrated an increasing trend for type I IFN (IFN-α2) relative to severe cases and healthy controls (Figure 4.12 A). In a separate study, IFN-α levels were also higher in asymptomatic patients relative to symptomatic COVID-19 patients.[241] The presence of type I IFN in moderate cases of COVID-19 in this study, along with a recent study by Lucas *et al.*, suggest that SARS-CoV-2 infection is capable of inducing a type I IFN response *in vivo*; however, emerging clinical data suggest that the extent and duration of type I IFN response may dictate the clinical course of COVID-19.[147,247] In this study, sera from COVID-19 patients were collected at admission (all < 21 days post symptom onset). Early induction of IFN-α2 in moderate cases of COVID-19 may provide an antiviral advantage. Indeed, this study was able to demonstrate that a range of physiologically detected concentrations of IFN-α2 was capable of inducing an antiviral response in human airway epithelial cells and that SARS-CoV-2 was unable to suppress ISG expression induced by physiologically detected concentrations of IFN-α2 (Figures 4.12 C-D). Furthermore, two different concentrations of IFN-α2 (1 and 10 ng/mL) were able

to reduce SARS-CoV-2 replication in human airway epithelial cells (Figure 4.12 E). Thus, extrapolating from *in vitro* testing of a range of IFN-α2 concentrations, it can be speculated that the levels of IFN-α2 detected in sera from moderate cases of COVID-19 would be sufficient to suppress SARS-CoV-2 replication. This study was unable to detect IFN-α2 in severe COVID-19 patients at the time of sample collection ($<$ 21 days from first symptom onset). Thus, early up-regulation of type I IFN responses such as IFN-α2 may be a predictor of moderate COVID-19 disease severity. Additional studies with later samples from severe COVID-19 patients will identify if there is a late and prolonged induction of type I IFNs as reported recently by Lucas *et al.*[247] In spite of recent progress in understanding type I IFN responses in COVID-19 patients, factors associated with early or delayed and short-acting versus prolonged type I IFN induction in COVID-19 patients are poorly understood. The *in vitro* experiments in this study are not sufficient to capture SARS-CoV-2-IFN interactions in a model of severe COVID-19. There is a need to develop appropriate animal models to accurately represent and study the full spectrum of COVID-19 disease severities.

Although recent studies have demonstrated the ability of SARS-CoV-2 to induce IFN responses,[314,430] other studies have demonstrated the ability of SARS-CoV-2 proteins to suppress IFN responses,[411] along with inducing a delayed type I IFN response in SARS-CoV-2-infected cells.[230] In this study, it is demonstrated that SARS-CoV-2 can induce a type I IFN response in human airway epithelial cells. The human respiratory tract is made up of more than one cell type that can be infected with SARS-CoV-2, thus it is important to characterize type I IFN responses in the full range of susceptible human airway and lung cell types. In this study, the ability of SARS-CoV-2 to mount a more potent IFN response in the absence of known IFN modulating viral proteins that have been identified in other studies was not assessed. Future studies will need to assess the full potential of IFN responses in cells infected with wild-type and deletion variants of SARS-CoV-2. More work is also needed to identify the detailed kinetics of IFN induction by SARS-CoV-2 RNA in human cells, followed by subsequent modulation of IFN responses by viral proteins. This will be particularly important to understand why some patients mount a detectable IFN response, whereas others do not. Timing, intensity, and duration of type I IFN responses will be important to understand the range of disease outcomes in COVID-19 patients. Other beta-coronaviruses continue to infect humans, along with infections with emerging variants of SARS-CoV-2. Thus, it is important to assess the efficacy of IFN responses against a range of human coronaviruses to determine differences in pathogenesis and disease severity.

In conclusion, this study demonstrates that SARS-CoV-2 is a weak stimulator of type I IFN production in infected human airway epithelial cells, relative to poly(I:C). However, the data suggest that low levels of type I IFN response in SARS-CoV-2-infected cells are sufficient to activate downstream expression of antiviral ISGs. In addition, the data in this

study demonstrate that SARS-CoV-2 is unable to inhibit downstream IFN responses that are mediated by STAT proteins, which is promising for the development of type I IFNs as treatment or post-exposure prophylactics.[167,297] Clinical trials for combination IFNβ therapy against MERS-CoV are currently ongoing.[23] IFNβ, in combination with lopinavir-ritonavir and ribavirin, has been used with promising results in COVID-19 patients.[182] Nebulized IFNβ is part of the standard of care for COVID-19 patients in China.[424] Furthermore, it is also demonstrated that levels of IFN-α2 detected in sera from patients with moderate COVID-19 can (A) induce an antiviral ISG response in human airway epithelial cells and (B) inhibit SARS-CoV-2 replication. Thus, this study highlights the dynamic nature of virus-host interaction during the course of SARS-CoV-2 infection and raises intriguing questions about the role and timing of IFN responses in predicting the likely severity of COVID-19.

4.3 Summary

The COVID-19 pandemic has emphasized the global efforts that are required to manage global problems. Researchers must join together to form interdisciplinary teams to develop strategies to effectively respond to current and future public health crises. Bioinformatics positions itself at the crux of such research teams, allowing for in-depth and efficient analysis of data, quick dissemination of that data, and the ability to aggregate, disaggregrate, and re-analyze previously existing data, all with the aim of answering pressing public health questions.

This data chapter has showcased how multi-omic approaches can be harnessed to uncover critical information related to both pathogen and host. Through a combination of transcriptomics, proteomics, and immunohistological validation, the research within chapter 4.1 challenges the current understanding of how the novel SARS-CoV-2 virus enters host cells. While the ACE2 receptor is still of significance to this viral entry mechanism, low levels of detection across the various methods reported in this chapter suggest that it may not be the entire story and further research is required to fully understand the means by which coronaviruses infect host cells. Using a similar combination of computational and wet-laboratory techniques, chapter 4.2 delves into the various host responses elicited once the SARS-CoV-2 virus has entered a cell which may provide insights into long-term outcomes and potential treatment methods.

The computational techniques outlined provide a blueprint for how bioinformatic analysis of large data cohorts can be leveraged to answer important biological questions in a timely manner. Devoting the necessary resources into sequencing analyses of large, diverse patient

cohorts (even after the pandemic officially ends) will ensure that when the next global disease outbreak occurs, scientists and public health officials and can more quickly and equitably respond.

Chapter 5

Conclusions

> [Scientists] are expanding their communication across fields [...] The result is the formation of new, interdisciplinary scientific fields, such as bioinformatics, that are painting expanded pictures of the world around us, thanks to the collective efforts of diverse groups of researchers.
>
> *The Benefits of Interdisciplinary Research: Our Experience With Pathogen Bioinformatics*
> Jennifer Gardy and Fiona Brinkman

Pulmonary diseases remain a major public health concern worldwide. Though there have been numerous advances in respiratory research over the past decades that have allowed for a deeper understanding of the biological mechanisms governing various lung diseases, early detection, sub-classification, and personalized treatment methods still elude us.[209] In recent years, the research and applications thereof made possible through bioinformatic analysis of large 'omic data sets have proven to be incredibly promising at filling in some of these knowledge gaps and pushing the field of respiratory medicine closer to its goals. In the following chapter, I briefly describe the contributions of the research presented in this thesis, discuss some of its limitations, and conclude with a prospective for the field.

5.1 Summary of major findings

5.1.1 Tobacco smoke exposure and viral mimic

It is universally understood that tobacco smoke exposure has myriad negative effects on the human body, particularly the lungs. Despite this fact, there are still an estimated 1.3 billion tobacco smokers worldwide[282] and there are still various aspects regarding the effects of tobacco smoke exposure on lung function, including damage and repair mechanisms, that remain poorly understood. It is therefore imperative that research continue on this topic to paint a fuller picture. As has been demonstrated in this thesis, bioinformatics provides a toolkit to help glean useful information that has been previously out of reach of gold standard methods. In Chapter 2.1, I demonstrate that by using bioinformatics to leverage previously deposited data sets, and by focusing on gene families that were previously understudied in the context of lung biology, such as the ABC transporters, new information regarding the effects of tobacco smoke exposure on lung function and disease onset can be obtained.

The analyses performed in Chapter 2.1[11] identify differential expression of multiple ABC transporters across patient cohorts. *ABCB6*, *ABCC1*, and *ABCC3*, for example, all show an increase in expression correlated with tobacco smoke exposure and, while expression of these genes do significantly decrease following smoking cessation, they do not return to the levels observed in the healthy, "never-smoker" cohort. As discussed in Chapter 2.1.4, *ABCB6* has been shown to elicit protective responses to reactive oxygen species,[219,290] specifically through the inhibition of carcinogenic cytotoxic components found in cigarette smoke.[286] It is therefore suggested that the observed up-regulation in *ABCB6* is likely an important initial protective response to cigarette smoke exposure, though potential consequences of long-term over-expression may exist related to reduced responsiveness to chemotherapeutic agents. *ABCC1* and *ABCC3* have also been shown to provide protective effects by transporting glutathione-conjugated anions from the cytosol to the extracellular compartment to prevent cellular accumulation of toxic metabolites[231,433] and, similar to *ABCB6*, *ABCC1*, and *ABCC3* are also capable of effluxing chemotherapeutics of broad function including anti-cancer functions. Collectively it is conceivable that *ABCB6*, *ABCC1*, and *ABCC3* gene expression levels are regulated by a feedback mechanism linked to oxidative stress and toxin exposure induced by cigarette smoke exposure to help facilitate anti-oxidant activities within airway epithelial cells.

Importantly, it was observed that as cigarette smokers progress to develop COPD, *ABCC1* gene expression in particular is elevated compared to individuals that do not have COPD. Linked to cigarette smoking, elevated *ABCC1* gene expression has been

observed in both non-small cell carcinoma and small-cell carcinoma lung cancers.[67] The contribution of *ABCC1* biology in response to cigarette smoke exposure and to the associated development of COPD and lung carcinoma is intriguing. An *in vitro* model of cigarette smoke extract conditioned media induction of *ABCC1* gene expression may be valuable in further interrogating this biology and the functional consequences in the context of COPD and lung carcinomas.

Respiratory viruses are another prevalent stimuli to which the human airway may be exposed and, just like tobacco smoke, there is still much that remains to be understood regarding how different viruses affect lung function and health. Again, we can turn to understudied genes such as *ABCF1* and through a combination of bioinformatics and functional knock-out studies, elucidate new correlations between this gene and viral defense mechanisms in the lung.

Under VACV-70 challenge, a model of dsDNA viral exposure, ABCF1 was linked to CXCL10 secretion and the gene pathways regulated by ABCF1 under VACV-70 challenge were associated with TLR signaling and intracellular signal transduction. Furthermore, Poly I:C, a dsRNA analog and TLR3 ligand, induced CXCL10 in an ABCF1 sensitive mechanism. Collectively, this study's data suggests that ABCF1 may regulate CXCL10 production downstream of dsDNA sensing mechanisms and TLR3 in human airway epithelial cells. It remains possible that ABCF1 can function to complement viral sensing mechanisms mediated by canonical dsRNA viral response machinery (e.g., RIG-I) as there are possible redundancies in viral sensing mechanisms in the cell.[274] The data from this chapter form the foundation to pursue precisely how ABCF1 is regulated and where it functions in the network of cytosolic nucleic acid sensors and immune responses in human airway epithelial cells.

The findings in Chapter 2, as well as previously identified causal relationships between ABC transporters and lung disease (e.g., ABCC7 and cystic fibrosis) emphasize the importance of ABC transporters in human lung health and highlight the necessity of leveraging bioinformatics and data-mining approaches to unearth these critical and previously overlooked connections.

5.1.2 Cannabis smoke exposure

Unlike tobacco smoke and respiratory viruses, which have well understood correlations to lung disease (even if some of the details remain to be explored, as mentioned above), the effects of cannabis smoke on lung function and whether there are any strong correlations between exposure and the onset of lung disease have yet to be established. Previous legal

restrictions on the use and consumption of cannabis have resulted in a lack of research into the area compared to tobacco smoke and almost no large-scale clinical trials or longitudinal studies have been performed to date. Chapter 2 of this thesis demonstrates how bioinformatics can be used to extract novel information from already well-studied exposure systems. In Chapter 3, bioinformatics is instead used to investigate the effects of stimuli with little prior data.

The *in vitro* smoke exposure studies performed in Chapters 3.1. and 3.2. underscore the similarities between tobacco and cannabis smoke exposure on epithelial barrier integrity, antiviral defense mechanisms, pro-inflammatory immune response, oxidative stress response, and response to long-acting β-agonist/glucocorticoid therapy. As evidenced by the degraded cell barrier integrity, decrease in anti-viral defense mechanisms, and increased inflammation and oxidative stress, cannabis smoke is not innocuous. It may be for these reasons that the use of vaporizors for cannabis consumption is increasing in Canada[69,70] as users may be looking to reduce some of the harmful effects of cannabis combustion that are being highlighted in the literature and news. However, research still needs to be performed to determine whether vaping, specifically of dried flower/bud (as cannabis oils have be linked to their own uniquely concerning respiratory outcomes and as the vaporization of dried cannabis would be the most reasonable alternative to combustion in terms of product similarity), actually mitigates these effects and should therefore be suggested as a safer user practice.

With the above in mind, we do know from previous research[376–378] that cannabis use confers some acute benefits to the lung and while reduced airway compliance and association with bronchitis and coughing have been noted,[374,379] cannabis smoking appears to lack the same direct, long-term correlation with certain lung diseases such as cancer that are found with tobacco smoke. A limitation of previous cannabis studies is the lack of mechanistic interrogation into the role that the endocannabinoid system contributes to the observed functional consequences of cannabis smoke, both protective and detrimental. Chapter 3.3. looks into the various cannabinoid receptors in the lung to provide a comprehensive profile of expression and correlation with factors such as sex and airway disease status.

To better understand the functional consequences of the endocannabinoid system in the epithelial cells of the human airway, I performed a characterization study using bioinformatic and complementary protein analysis approaches. *In situ* and *in vitro* protein analysis demonstrated that three cannabinoid receptors, CB1, CB2, and TRPV1, are expressed at the protein level in human airway epithelium. The presence of these receptors provided the initial confirmation required for an exploration into the endocannabinoid system downstream of the receptors. Using 1090 unique patient samples of airway epithelial cells curated from publicly available data sets, I demonstrated in healthy subjects that

the gene expression levels of the endocannabinoid system are elevated in females relative to males. However, further analysis of patients with chronic lung diseases such as COPD and asthma suggest that disease status appears to be a stronger driver of endocannabinoid system gene expression, as the sex-dependent effects identified in healthy individuals are lost in individuals with these diseases. Lastly, my bioinformatic approach was validated by demonstrating that TRPV1, a top candidate up-regulated at the gene level in these studies, is also up-regulated at the protein level in asthmatics. Collectively, these results confirm the expression of the endocannabinoid system in human airway epithelial cells and suggest that both sex and disease status may impact cannabinoid responses. These results and those in the literature suggest that a universal response to inhaled cannabis by healthy subjects and those with asthma or COPD should not be assumed and cautions translation of safety and efficacy studies performed in healthy individuals to those with underlying asthma or COPD. Adjusting the response to cannabis use is especially pertinent with respect to sex and gender, as cannabis use patterns have been reported to differ among males and females, with females consuming less overall quantity and frequency.[68–70] Crucial sex differences identified in both the literature and in the surveyed user responses above mean that cannabis may affects males and females differently, suggesting there is no "one size fits all" strategy for the regulation of cannabis or for suggested user practice guidelines. As will be discussed later in this chapter, these results emphasize the importance of disaggregating data based on key phenotypic groups such as sex, age, and ethnicity to ensure the bioinformatic analyses performed are applicable to more than just white, cis-male bodies, which can often be considered the "default" if sex/gender and ethnic considerations are not taken into account.[298]

5.1.3 SARS-CoV-2 infection

The first two data chapters in this thesis demonstrated how bioinformatics can be used to glean new information regarding the effects of relatively well-studied airway stimuli (Chapter 2) and how similar computational methods can be leveraged to characterize the effects of under-studied stimuli (Chapter 3). Chapter 4 of this thesis uses bioinformatics to gain insight into a completely novel respiratory stimuli, completing our journey from the well-known to completely unknown. While respiratory viruses are broadly understood, as discussed in Chapter 2.2., the SARS-CoV-2 coronavirus did not exist until late 2019, requiring a vast array of multidisciplinary researchers to come together to characterize both the mechanisms and host outcomes of this novel pathogen to combat the resulting deadly disease pandemic.

The seminal report identifying angiotensin-converting enzyme 2 (ACE2) as the primary

receptor for SARS-CoV-1 did so through co-immunoprecipitation of ACE2 with the SARS-CoV-1 spike (S) domain 1 in a HEK293 cell over-expression system.[238] ACE2, in conjunction with the cellular serine protease TMPRSS2, were validated as cellular entry determinants for SARS-CoV-1 using mechanistic studies.[256] These findings laid the foundation for our understanding of how coronaviruses enter their target cells via the binding of a surface unit of the S glycoprotein to a cellular receptor followed by protease facilitated cleavage of the S protein leading to fusion of the viral and cellular membranes.

Following the initial outbreak of SARS-CoV-2 in 2019, studies began to suggest that this novel coronavirus employed both the ACE2 receptor and cellular proteases such as TMPRSS2, CTSB, and CTSL for host cell entry in a similar manner to SARS-CoV-1.[169] The binding affinity between ACE2 and the SARS-CoV-2 S protein were suggested to be sufficient for human transmission[341] and ACE2 was subsequently detected in human alveoli,[149] airway epithelial cells,[149] A549 type II alveolar epithelial cells,[149] and human lung adenocarcinoma cell line Calu-3[385] via *in situ* immunohistochemical and immunocytochemical staining. TMPRSS2 was also observed to be expressed in the human airway epithelium and type II alveolar epithelial cells.[46]

While these studies provide an important starting point for our understanding of coronavirus cellular entry mechanisms, certain methodological issues need to be addressed. For instance, in the original report suggesting ACE2 as the primary receptor for SARS-CoV-2 only one antibody was used and the specificity of this antibody was overlooked.[149] A rigorous protein profiling study undertaken by researchers at the Human Protein Atlas observed strong staining for ACE2 in several organs that lacked mRNA expression when using this original antibody, suggesting a lack of binding specificity.[161] Indeed, the immunohistochemical analyses performed in Chapter 4.1 found very limited expression of ACE2 in the respiratory system using this same antibody.[12] Therefore, the specificity of the ACE2 antibody used for analysis of expression patterns in human lung tissues remains to be addressed. Another issue with the seminal studies regarding the SARS-CoV-2::ACE2 association is that there is no comparison of ACE2 expression levels in the lung with those observed in any other tissues or organs in the human body. The immunohistochemical analysis performed by the Human Protein Atlas not only used two different and stringently validated antibodies, but also utilized an extended patient cohort with samples from 61 different organs and tissue types including 360 different normal lung samples analyzed by three independent consortia (Human Protein Atlas, FANTOM5, and GTEx).[161] This analysis shows very little ACE2 expression in the human airway. Rather, the highest levels of expression of ACE2 occur in the gastrointestinal tract, kidneys, and testis. While these results do not necessarily negate the potential importance of the ACE2 receptor in relation to SARS-CoV-2 entry mechanisms, it certainly warrants further investigation.

To address the uncertainties related to SARS-CoV-2 receptors in the human lung, gene expression and *in situ* protein profiling of candidate receptors in human airway epithelial cells and lung tissue were performed in Chapter 4.1. At both the transcriptomic (microarray, RNA-seq, and scRNA-seq) and proteomic (Human Proteome Map and immunohistochemical analysis) levels, ACE2 expression in lung epithelial cells throughout the airway tree was either not detected or detected at marginal levels. Analysis of the promoter regions of putative SARS-CoV-2 receptors also suggested low ACE2 expression in the lung. These results do not appear to be correlated with key phenotypes such as sex and age.

Presence of other candidate receptors and/or proteases such as TMPRSS2, CD147, and GRP78 were confirmed via *in situ* protein expression analysis, with CD147 and GRP78 in particular found to be highly expressed across all methods of analysis, from the trachea down to the epithelium of the small airways. These findings suggest that while ACE2 may still be an important receptor for the S protein of SARS-CoV-2 and other similar coronaviruses, with countless studies identifying it as integral for infection, other mechanisms of viral entry are likely at play and warrant further investigation. Recent studies have proposed that SARS-CoV-2 may first enter through the eye conjunctiva and cells in the upper airways were ACE2 expression is higher whereby up-regulation of ACE2 expression through interferon stimulation or external factors such as tobacco smoking allows the virus to spread and infect the lung parenchyma.[161] Once again, these putative mechanisms require further validation.

Regardless of the method of entry, it is also important to understand what happens once the virus has breached the barrier defenses of it's host. Chapter 4.2 interrogates the human host response upon viral entry and gains important insights into type I interferons and interferon-stimulating genes (ISGs), both critical components of the host defense response. The time-series RNA-seq study in Chapter 4.2 identifies a weak stimulation of type I interferon transcript expression over time, correlating with an increase in viral replication. These results demonstrate that infection with wild-type and replication competent SARS-CoV-2 is capable of inducing a type I IFN response in human airway epithelial cells sufficient to activate downstream anti-viral ISGs. In addition, the data in this study demonstrate that SARS-CoV-2 is unable to inhibit downstream IFN responses that are mediated by STAT proteins, which is promising for the development of type I IFNs as treatment or post-exposure prophylactics.[167,297]

Furthermore, it is also demonstrated that levels of IFN-α2 detected in sera from patients with moderate COVID-19 can induce an antiviral ISG response in human airway epithelial cells as well as inhibit SARS-CoV-2 replication. Thus, this study highlights the dynamic nature of virus-host interaction during the course of SARS-CoV-2 infection and raises intriguing questions about the role and timing of IFN responses in predicting the likely severity of COVID-19.

Another key take-away from Chapter 4 is the observed benefit of multi-omic analysis through the combination of transcriptomics (both bulk and single cell) and proteomics, specifically in Chapter 4.1. The multi-disciplinary approach of combining this type of large, multi-omic analysis with *in situ* validation techniques is something that will be discussed further on in this chapter as paramount for the future of respiratory medicine.

It is unlikely that the COVID-19 pandemic will be the last global disease outbreak experienced in our lifetime. As climate change continues to destabilize our ecological surroundings, and humans continue to remove or encroach upon the habitats of wild animals and engage in controversial farming practices, zoonotic viruses will continue to be of major concern to the global health sector. Our ability to protect ourselves from emerging pathogenic organisms relies heavily on our ability to combine multidisciplinary skill-sets to understand the mechanisms and consequences of previous disease outbreaks. Bioinformatics coupled with large-scale genomic surveillance efforts around the world offers a toolkit to apply what we have learned to the next global health crisis. It is imperative to learn how to use this toolkit effectively so that future responses are swift and equitable.

5.2 Limitations and future directions of bioinformatic analysis

It is evident that high-throughput sequencing, particularly of RNA, and the various downstream computational work flows have become critical in the advancement of respiratory medicine. However, there are still areas that require improvement as this type of research becomes the new standard. My aim is to shed light on these limitations and provide suggestions on how we as a scientific community can continue to develop bioinformatic methodologies for the benefit of both basic science and clinical health research.

One of the hurdles to bioinformatic analysis of respiratory diseases is access to sufficiently large and high quality data sets. On the one hand, some respiratory diseases are relatively rare and, as such, studies can often be under-powered due to small cohort size.[156] On the other hand, some of the most common, and often chronic, respiratory diseases, such as COPD and asthma, are incredibly complex and involve multiple pathogenetic mechanisms resulting in diverse clinical expressions all warranting in-depth analysis. Experiments can be performed *in vitro* using cell lines to gain a preliminary understanding of the effects a specific change may have on the airway epithelium (as is demonstrated in Chapters 3.1 and 3.2 of this thesis) however, primary samples from human subjects are ultimately considered the best practice, especially when studying the intersection of external stimuli

and preexisting respiratory conditions. A coordinated effort to increase data-banking practices in repositories such as the Gene Expression Omnibus (GEO) will allow for the sharing of data and the potential to increase population sizes in some studies. Chapters 2.1 and 4.1 of this thesis in particular demonstrate the important and novel findings that can be obtained through the curation and analysis of previously deposited data sets. However, a key caveat of the success of this type of data deposition and subsequent analysis is the type and quality of data that is being collected and shared. As described in Chapters 2.1.4 and 4.1.4 of this thesis, the curation and downstream analysis of data sets from different studies and research labs requires a considered approach with respect to data normalization and correction of batch effects, as deviations in sample collection and processing protocols can have major impacts.[195,229] This type of data pre-processing is critical for drawing any conclusions as many of these data sets would not be comparable otherwise.

Microarray data can pose particular challenges, especially when it comes to batch correction and data processing. Microarray technology has been the gold standard for transcriptomic analysis for years. Since microarrays have been widely used by many research groups around the world there is both a large repository of data available, as mentioned above, as well as a thorough understanding of the limitations of the technology. For transcriptomic analysis, microarrays have proven to be an inexpensive and readily accessible option for biomarker discovery and detection. Microarray expression profiling of epithelial brushings in asthma was successfully used to stratify patients into distinct subtypes of mild to moderate asthma with a Th2-high or Th2-low phenotype.[401] However, microarrays suffer from low signal-to-noise ratio and are limited to the detection of transcripts for which a probe exists on the given microarray chip, both limiting the ability of detecting lowly abundant or novel transcripts in a given sample.[401] Furthermore, the number of microarray platforms on the market is quite extensive and, with no two microarray chips being created equal, the ability to compare results across platforms, which is paramount for the successful leveraging of deposited data sets, as mentioned above, requires stringent correction methods and probe annotation. As sequencing of RNA has become more feasible in recent years, it has been successfully applied to fill into some of the gaps left by microarray technology. RNA-seq is a high-throughput, dynamic alternative to microarrays for transcriptomic analysis that is highly sensitive to lowly expressed transcripts and does not restrict the analysis to known transcripts, providing a path for unbiased, exploratory data analysis and biomarker discovery.

Though the benefits of transitioning from microarray to RNA-seq technology are hopefully apparent, especially in instances of whole transcriptomic or exploratory analyses, some considerations also need to be made to ensure the best results. Sequencing depth needs to be considered when designing a transcriptomic analysis, as low depth and coverage can

lead to poor recall which can result in an increase in false negatives and an inability to detect more rare variants. 30x sequencing depth is considered good for most studies and should be considered early in the experimental design process.[156] The RNA-seq studies outlined in this thesis have a sequencing depth of 30 million reads/sample, providing sufficient depth and coverage of the transcriptome. RNA-seq also places an extensive demand on bioinformatics and statistics for storage, processing analysis, and interpretation; far greater than that of it's microarray counterpart. The average .FASTQ file output from an RNA-seq run can range from 5-25 GB compared to the average .CEL file out from an Affymetrix microarray analysis which is approximately 50 MB. As sequencing depth and sample size increase, both necessary for sufficient coverage and statistical power, the storage and processing capabilities required for an RNA-seq analysis grow exponentially. As such, appropriate resources, both structural and financial, need to be allotted for 'omics analyses of large biological data sets.

Another major consideration for effective computational analysis of respiratory data is the need for consistent and thorough patient phenotyping and a concerted effort to accumulate this type of detailed metadata. Sex, gender, ethnicity, age, and co-morbidity are all important co-factors that can affect both the diagnosis and prognosis of respiratory illness. A good example of this is with idiopathic pulmonary fibrosis (IPF), a chronic lung disease of unknown etiology that generally affects adults over the age of 50. Patient gender has been identified as an influencing factor with respect to the diagnosis of IPF, with women possibly under-diagnosed and men over-diagnosed with the disease, especially when exam results are inconclusive.[26] However, male sex has also been independently associated with an increased risk of poorer prognostic outcomes such as lung transplantation and death after adjusting for age and lung function metrics.[432] The disaggregation of data by sex and gender is critical for this type of understanding and is necessary moving forward. As described in Chapter 3.3 of this thesis, the direct comparison of endocannabinoid system expression in males versus females identified important sex differences that, in the context of gendered cannabis use data, could lead to considerably different outcomes and require targeted approaches between males and females in terms of cannabis-related disease prevention. Chapter 3 also highlights the importance of disclosing relevant patient co-morbidities such as lung disease and smoking status, especially in the context of evolving research areas such as cannabis exposure. Previous criminalization of cannabis use has made research on the topic exceedingly difficult. Many of the existing studies on cannabis exposure suffer from a lack of description of consumption method and furthermore, the tobacco exposure history for many individuals identified as cannabis users is unclear. With the legalization of cannabis in Canada, researchers are now able to begin studies into acute and chronic cannabis use but, to do so, will need sufficiently sized cohorts of individuals who only smoke

cannabis so as to accurately characterize the short- and long-term effects of it's use. Efforts were also made to separate patient data by sex and age in Chapter 4.1 to ascertain whether there were any significant differences in SARS-CoV-2 receptor expression. However, as described in Chapter 4.1.4, not all data sets analyzed provided the necessary metadata for robust disaggregation based on sex and age, and even fewer provided detailed phenotyping of ethnic background. Moving forward, it is imperative that clinicians, researchers, and data banks work together to develop and follow guidelines for the collection, reporting, and use of phenotypic data. Bioinformatics is only as powerful as the data collected. We need to be harnessing the full potential of computational biology while also ensuring all research outcomes are more broadly applicable.

With respect to the future of RNA-seq in respirology, a strong push should be made for the development and funding of more longitudinal studies. RNA-seq inherently provides a snap shot of what is going on within a cell at a given time. While this is useful for providing foundational knowledge on the affects of a given airway stimuli or for assessing acute respiratory conditions, many of the most common airway diseases are both chronic and complex and therefore cannot be sufficiently characterized by a single sequencing experiment. To fully understand the impacts environmental exposures such as smoke or respiratory viruses have on the human lung, and what the relationship is to disease onset and progression, large cohorts of well-phenotyped patients will need to be followed over longer periods of time, with sampling frequency being dependent on study design, exposure type, and expected disease outcome. Considerations will also need to be taken into which compartments are sampled overtime, as epithelial cells from the lower airways are not readily accessible for repeated sampling compared to cells derived from the nasopharyngeal airways. The use of multi-omic technologies to identify surrogate signals of lung pathology circulating in the blood or saliva would be of further benefit for the above proposed longitudinal studies. While the cost of sequencing has indeed decreased dramatically over the last decade, this type of proposed large-scale longitudinal sequencing study is still quite expensive. However, if we are to effectively characterize the mechanisms and outcomes of these complex respiratory diseases on our path towards precision medicine, these types of studies are absolutely necessary.

The final key direction in which I believe both respiratory biology and general science need to head is towards the further application of multi-omic analyses, particularly in clinical research. Multi-omic approaches are critical for future medical advances such as the development of prognostic biomarkers or novel drug candidates[216] and will require sufficient funding and skilled research analysts to translate the large amounts of data into medically relevant outcomes.

As noted previously, this thesis has primarily focused on a single 'omic technology:

RNA-seq. While the transcriptomic analyses presented in this thesis have led to important and novel insights into respiratory biology, there are certain biologically relevant questions that assessment of RNA alone cannot answer. For example, while RNA abundance is a reasonably good indicator of function and often correlates well with downstream protein abundance, proteomic validation would no doubt be incredibly useful and perhaps further elucidate the disease mechanisms being studied. As demonstrated in Chapter 4.1, the combination and comparison of bulk RNA-seq, scRNA-seq, and proteomics provides deeper insights while simultaneously validating previous findings thus, strengthening the research as a whole and improving scientific rigour. Single cell sequencing, for example, provides a more granular understanding of cellular responses to stimuli which is especially important given the heterogeneity found between individual cells and cell types in chronic, complex lung diseases. The combination of bulk and single cell RNA-seq may lead to even more precise and rapid detection of aberrant function and is just one example of the benefits of integrative 'omics.

Multi-omic analysis is not without it's challenges however. The inherent differences between different 'omic data sets in terms of size and retrievable output makes data harmonization (e.g., normalization, batch correction, and other transformations) both vitally important and incredibly complex. Data heterogeneity is often one of the primary bottlenecks for multi-omic workflows and requires considered pre-processing to overcome. Some tools have been developed in recent years to improve handling of multi-omic data sets, such as Multi-Omics Factor Analysis (MOFA), but an understanding of the variability between data sets is still necessary.[216] Furthermore, differences in information burden (i.e., the amount of data generated between each technique) may lead to information bias, as the large amounts of output from one 'omic data set may overshadow that of another.[216] For example, transcriptomic data sets are comprised of hundreds of thousands of transcripts compared to the few thousand proteins detectable in a proteomic data set, making interpretation a challenge. Finally, researchers will also need to be cautious when choosing the statistical tools and models used to analyze their multiple data sets to ensure data is not over-simplified and to ensure robust and reproducible results. Machine-learning/deep-learning algorithms and statistical bench-marking studies are currently attempting to address some of these issues however, the fields are still in their infancy compared to gold-standard biological methods and more research is warranted.

Given sufficient funding and the expertise of bioinformaticians and other interdisciplinary researchers, the field of multi-omics will continue to advance us towards a future of precision or even personalized medicine. The ultimate purpose of applied multi-omics is to increase the diagnostic yield for health and improve disease prognosis via robust understanding of genotype-to-phenotype relationship.

5.3 Final remarks

In spite of the aforementioned limitations, a few key conclusions can be gleaned from this thesis. While the large amounts of data can be a challenge, it is also an exciting resource.

Success in precision medicine relies heavily on high-quality genetic and molecular data sets derived from large, well-phenotyped patient cohorts.[133] Bioinformaticians offer the statistical and computational expertise needed to integrate large amounts of multi-omic data and make significant meaning from it. It is through these advances that the medical field inches closer to the goal of personalized medicine for all.

In the broad sense, the future of bioinformatics is definitely bright. The potential advantages that come from combining both population and individual patient data with epidemiological and molecular profiling are innumerable, ranging from improved disease prevention, more accurate diagnosis, tailored treatment and improved long term outcomes for a variety of human health conditions.[133] As we focus in on respiratory medicine in particular, bioinformatic analysis of large mutli-omic data will be of vital importance in tackling some of the greatest threats facing global respiratory health. As previously covered in this thesis, the progress made leading up to and during the COVID-19 pandemic will become even more critical as more novel respiratory viruses emerge in the coming years. Coupled with the already observable effects of climate change on respiratory health such as increased pollution and decreasing air quality exacerbating pre-existing health conditions such as asthma,[109] the need to be able to rapidly and accurately analyze diverse biological data sets for the benefit of both treatment and prevention has hit a new level of urgency, requiring a multidisciplinary approach with bioinformatics at the forefront.

The development of bioinformatic and 'omics research is happening at a rapid rate and while there are some 'omics methods that were unable to be fully discussed within this thesis, it is clear that, even while focusing primarily on RNA, NGS and large amounts of 'omics data combined with advances in bioinformatic tools and techniques have provided a wealth of knowledge and new opportunities for biologists, including those studying respiratory medicine.

As researchers begin to combine these methods and adapt their experimental designs to incorporate multi-omics integration, our collective understanding of the human airway will increase dramatically and result in more prevention and treatment of lung disease and improvement in the quality of life for many.

Letter of Copyright Permission

The published articles included in this thesis are made with the permission of all publishers. At the time of writing (October 2021), the policies of the publishers Nature, Frontiers, Wiley, The European Respiratory Society, and Cell do not require additional permissions for the reproduction of articles in theses. The relevant policies can be found at the following links:

1. Nature

   www.nature.com/nature-research/reprints-and-permissions/permissions-requests

2. Frontiers

   www.frontiersin.org/about/policies-and-publication-ethics

3. Wiley

   www.physoc.onlinelibrary.wiley.com/hub/permissions

4. The European Respiratory Society

   www.erj.ersjournals.com/authors/copyright

5. Cell Press

   www.cell.com/trends/editorial-policies

References

[1] Common asthma triggers. Report, Centers for Disease Control and Prevention, 2010.

[2] Global burden of disease. Database, Institute of Health Metrics, 2019.

[3] Middle east respiratory syndrome coronavirus (mers-cov). Report, World Health Organization, 2019.

[4] What are the risk factors for lung cancer? Report, Centers for Disease Control and Prevention, 2020.

[5] S. J. Abdallah, B. M. Smith, M. A. Ware, M. Moore, P. Z. Li, J. Bourbeau, and D. Jensen. Effect of vaporized cannabis on exertional breathlessness and exercise endurance in advanced chronic obstructive pulmonary disease. a randomized controlled trial. *Annals of the American Thoracic Society*, 15(10):1146–1158, 2018.

[6] D. I. Abrams. The therapeutic effects of cannabis and cannabinoids: An update from the national academies of sciences, engineering and medicine report. *European Journal of Internal Medicine*, 49:7–11, 2018.

[7] I. Abugessaisa, S. Noguchi, A. Hasegawa, J. Harshbarger, A. Kondo, M. Lizio, J. Severin, P. Carninci, H. Kawaji, and T. Kasukawa. FANTOM5 CAGE profiles of human and mouse reprocessed for GRCh38 and GRCm38 genome assemblies. *Scientific Data*, 4(1), 2017.

[8] D. Acharya, G. Liu, and M. U. Gack. Dysregulation of type i interferon responses in COVID-19. *Nature Reviews Immunology*, 20(7):397–398, 2020.

[9] A. Agrawal and M. T. Lynskey. Candidate genes for cannabis use disorders: findings, challenges and directions. *Addiction*, 104(4):518–532, 2009.

[10] J. A. Aguiar, R. D. Huff, W. Tse, M. R. Stämpfli, B. J. McConkey, A. C. Doxey, and J. A. Hirota. Transcriptomic and barrier responses of human airway epithelial cells exposed to cannabis smoke. *Physiological reports*, 7(20):e14249, 2019.

[11] J. A. Aguiar, A. Tamminga, B. Lobb, R. D. Huff, J. P. Nguyen, Y. Kim, A. Dvorkin-Gheva, M. R. Stämpfli, A. C. Doxey, and J. A. Hirota. The impact of cigarette smoke exposure, copd, or asthma status on abc transporter gene expression in human airway epithelial cells. *Scientific reports*, 9(1):1–12, 2019.

[12] J. A. Aguiar, B. J. Tremblay, M. J. Mansfield, O. Woody, B. Lobb, A. Banerjee, A. Chandiramohan, N. Tiessen, Q. Cao, A. Dvorkin-Gheva, et al. Gene expression and in situ protein profiling of candidate sars-cov-2 receptors in human airway epithelial cells and lung tissue. *European Respiratory Journal*, 56(3), 2020.

[13] A. Agustí, J. M. Antó, C. Auffray, F. Barbé, E. Barreiro, J. Dorca, J. Escarrabill, R. Faner, L. I. Furlong, J. Garcia-Aymerich, et al. Personalized respiratory medicine: exploring the horizon, addressing the issues. summary of a brn-ajrccm workshop held in barcelona on june 12, 2014. *American journal of respiratory and critical care medicine*, 191(4):391–401, 2015.

[14] S. Ahmadi, Z. Bozoky, M. D. Paola, S. Xia, C. Li, A. P. Wong, L. Wellhauser, S. V. Molinski, W. Ip, H. Ouyang, J. Avolio, J. D. Forman-Kay, F. Ratjen, J. A. Hirota, J. Rommens, J. Rossant, T. Gonska, T. J. Moraes, and C. E. Bear. Phenotypic profiling of CFTR modulators in patient-derived respiratory epithelia. *npj Genomic Medicine*, 2(1), 2017.

[15] M. O. Aksoy, V. Kim, W. D. Cornwell, T. J. Rogers, B. Kosmider, K. Bahmed, C. Barrero, S. Merali, N. Shetty, and S. G. Kelsen. Secretion of the endoplasmic reticulum stress protein, GRP78, into the BALF is increased in cigarette smokers. *Respiratory Research*, 18(1), 2017.

[16] A. A. Al-Hashimi, P. Lebeau, F. Majeed, E. Polena, Š. Lhotak, C. A. Collins, J. H. Pinthus, M. Gonzalez-Gronow, J. Hoogenes, S. V. Pizzo, M. Crowther, A. Kapoor, J. Rak, G. Gyulay, S. D'Angelo, S. Marchiò, R. Pasqualini, W. Arap, B. Shayegan, and R. C. Austin. Autoantibodies against the cell surface–associated chaperone GRP78 stimulate tumor growth via tissue factor. *Journal of Biological Chemistry*, 292(51):21180–21192, 2017.

[17] S. Aldington, M. Williams, M. Nowitz, M. Weatherall, A. Pritchard, A. McNaughton, G. Robinson, and R. Beasley. Effects of cannabis on pulmonary structure, function and symptoms. *Thorax*, 62(12):1058–1063, 2007.

[18] A. F. Altenburg, S. E. van Trierum, E. de Bruin, D. de Meulder, C. E. van de Sandt, F. R. M. van der Klis, R. A. M. Fouchier, M. P. G. Koopmans, G. F. Rimmelzwaan, and R. D. de Vries. Effects of pre-existing orthopoxvirus-specific immunity on the performance of modified vaccinia virus ankara-based influenza vaccines. *Scientific Reports*, 8(1), 2018.

[19] G. D. Amatngalim, W. Broekman, N. M. Daniel, L. E. P. M. van der Vlugt, A. van Schadewijk, C. Taube, and P. S. Hiemstra. Cigarette smoke modulates repair and innate immunity following injury to airway epithelial cells. *PLOS ONE*, 11(11):e0166255, 2016.

[20] G. D. Amatngalim, J. A. Schrumpf, A. Henic, E. Dronkers, R. M. Verhoosel, S. R. Ordonez, H. P. Haagsman, M. E. Fuentes, S. Sridhar, J. Aarbiou, R. A. Janssen, A. N. Lekkerkerker, and P. S. Hiemstra. Antibacterial defense of human airway epithelial cells from chronic obstructive pulmonary disease patients induced by acute exposure to nontypeable haemophilus influenzae: Modulation by cigarette smoke. *Journal of Innate Immunity*, 9(4):359–374, 2017.

[21] Z. Ammous, N. R. Hackett, M. W. Butler, T. Raman, I. Dolgalev, T. P. O'Connor, B.-G. Harvey, and R. G. Crystal. Variability in small airway epithelial gene expression among normal smokers. *Chest*, 133(6):1344–1353, 2008.

[22] J. Andrae, R. Gallini, and C. Betsholtz. Role of platelet-derived growth factors in physiology and medicine. *Genes & Development*, 22(10):1276–1312, 2008.

[23] Y. M. Arabi, , A. Y. Asiri, A. M. Assiri, H. A. A. Jokhdar, A. Alothman, H. H. Balkhy, S. AlJohani, S. A. Harbi, S. Kojan, M. A. Jeraisy, A. M. Deeb, Z. A. Memish, S. Ghazal, S. A. Faraj, F. Al-Hameed, A. AlSaedi, Y. Mandourah, G. A. A. Mekhlafi, N. M. Sherbeeni, F. E. Elzein, A. Almotairi, A. A. Bshabshe, A. Kharaba, J. Jose, A. A. Harthy, M. A. Sulaiman, A. Mady, R. A. Fowler, F. G. Hayden, A. Al-Dawood, M. Abdelzaher, W. Bajhmom, and M. A. Hussein. Treatment of middle east respiratory syndrome with a combination of lopinavir/ritonavir and interferon-β1b (MIRACLE trial): statistical analysis plan for a recursive two-stage group sequential randomized controlled trial. *Trials*, 21(1), 2020.

[24] T. M. Araújo, A. D. Seabra, E. M. Lima, P. P. Assumpção, R. C. Montenegro, S. Demachki, R. M. Burbano, and A. S. Khayat. Recurrent amplification of RTEL1 and ABCA13 and its synergistic effect associated with clinicopathological data of gastric adenocarcinoma. *Molecular Cytogenetics*, 9(1), 2016.

[25] H. Arora, S. M. Wilcox, L. A. Johnson, L. Munro, B. A. Eyford, C. G. Pfeifer, I. Welch, and W. A. Jefferies. The ATP-binding cassette gene ABCF1 functions as an e2 ubiquitin-

conjugating enzyme controlling macrophage polarization to dampen lethal septic shock. *Immunity*, 50(2):418–431.e6, 2019.

[26] D. Assayag, J. Morisset, K. A. Johannson, A. U. Wells, and S. L. Walsh. Patient gender bias on the diagnosis of idiopathic pulmonary fibrosis. *Thorax*, 75(5):407–412, 2020.

[27] O. Auerbach, E. C. Hammond, and L. Garfinkel. Changes in bronchial epithelium in relation to cigarette smoking, 1955–1960 vs. 1970–1977. *New England Journal of Medicine*, 300(8):381–386, 1979.

[28] M. Aufderheide, S. Scheffler, S. Ito, S. Ishikawa, and M. Emura. Ciliatoxicity in human primary bronchiolar epithelial cells after repeated exposure at the air–liquid interface with native mainstream smoke of k3r4f cigarettes with and without charcoal filter. *Experimental and Toxicologic Pathology*, 67(7-8):407–411, 2015.

[29] T. Ayakannu, A. H. Taylor, T. H. Marczylo, J. M. Willets, and J. C. Konje. The endocannabinoid system and sex steroid hormone-dependent cancers. *International Journal of Endocrinology*, 2013:1–14, 2013.

[30] K. L. Bailey, D. J. Romberger, D. M. Katafiasz, A. J. Heires, J. H. Sisson, T. A. Wyatt, and E. L. Burnham. Tlr 2 and tlr 4 expression and inflammatory cytokines are altered in the airway epithelium of those with alcohol use disorders. *Alcoholism: Clinical and Experimental Research*, 39(9):1691–1697, 2015.

[31] K. L. Bailey, T. A. Wyatt, D. M. Katafiasz, K. W. Taylor, A. J. Heires, J. H. Sisson, D. J. Romberger, and E. L. Burnham. Alcohol and cannabis use alter pulmonary innate immunity. *Alcohol*, 80:131–138, 2019.

[32] G. BALDWIN, D. TASHKIN, D. BUCKLEY, A. PARK, S. DUBINETT, and M. ROTH. Marijuana and cocaine impair alveolar macrophage function and cytokine production. *American Journal of Respiratory and Critical Care Medicine*, 156(5):1606–1613, 1997.

[33] A. Banerjee, K. Baid, and K. Mossman. Molecular pathogenesis of middle east respiratory syndrome (MERS) coronavirus. *Current Clinical Microbiology Reports*, 6(3):139–147, 2019.

[34] A. Banerjee, N. El-Sayes, P. Budylowski, D. Richard, H. Maan, J. Aguiar, K. Baid, M. R. D'Agostino, J. C. Ang, B. J.-M. Tremblay, et al. Experimental and natural evidence of sars-cov-2 infection-induced activation of type i interferon responses. *iScience*, 24(5):102477, 2021.

[35] A. Banerjee, J. A. Nasir, P. Budylowski, L. Yip, P. Aftanas, N. Christie, A. Ghalami, K. Baid, A. R. Raphenya, J. A. Hirota, M. S. Miller, A. J. McGeer, M. Ostrowski, R. A. Kozak, A. G. McArthur, K. Mossman, and S. Mubareka. Isolation, sequence, infectivity, and replication kinetics of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*, 26(9):2054–2063, 2020.

[36] A. Banerjee, N. Rapin, T. Bollinger, and V. Misra. Lack of inflammatory gene expression in bats: a unique role for a transcription repressor. *Scientific Reports*, 7(1), 2017.

[37] A. Banerjee, X. Zhang, A. Yip, K. S. Schulz, A. T. Irving, D. Bowdish, B. Golding, L.-F. Wang, and K. Mossman. Positive selection of a serine residue in bat IRF3 confers enhanced antiviral protection. *iScience*, 23(3):100958, 2020.

[38] A. J. Barton, J. Hill, A. J. Pollard, and C. J. Blohmke. Transcriptomics in human challenge models. *Frontiers in immunology*, 8:1839, 2017.

[39] P. Bastard, L. B. Rosen, Q. Zhang, E. Michailidis, H.-H. Hoffmann, Y. Zhang, K. Dorgham, Q. Philippot, J. Rosain, V. Béziat, et al. Autoantibodies against type i ifns in patients with life-threatening covid-19. *Science*, 370(6515), 2020.

[40] S. R. Bates, J.-Q. Tao, H. L. Collins, O. L. Francone, and G. H. Rothblat. Pulmonary abnormalities due to ABCA1 deficiency in mice. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 289(6):L980–L989, 2005.

[41] C. T. Bauch, J. O. Lloyd-Smith, M. P. Coffee, and A. P. Galvani. Dynamically modeling SARS and other newly emerging respiratory illnesses. *Epidemiology*, 16(6):791–801, 2005.

[42] C. M. Bauer, S. J. DeWitte-Orr, K. R. Hornby, C. C. Zavitz, B. D. Lichty, M. R. Stämpfli, and K. L. Mossman. Cigarette smoke suppresses type i interferon-mediated antiviral immunity in lung fibroblast and epithelial cells. *Journal of Interferon & Cytokine Research*, 28(3):167–179, 2008.

[43] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[44] J. M. Bergelson. Isolation of a common receptor for coxsackie b viruses and adenoviruses 2 and 5. *Science*, 275(5304):1320–1323, 1997.

[45] D. Bernhard, C. W. Huck, T. Jakschitz, G. Pfister, B. Henderson, G. K. Bonn, and G. Wick. Development and evaluation of an in vitro model for the analysis of cigarette smoke effects on cultured cells and tissues. *Journal of Pharmacological and Toxicological Methods*, 50(1):45–51, 2004.

[46] S. Bertram, A. Heurich, H. Lavender, S. Gierer, S. Danisch, P. Perin, J. M. Lucas, P. S. Nelson, S. Pöhlmann, and E. J. Soilleux. Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS ONE*, 7(4):e35876, 2012.

[47] V. Besnard, Y. Matsuzaki, J. Clark, Y. Xu, S. E. Wert, M. Ikegami, M. T. Stahlman, T. E. Weaver, A. N. Hunt, A. D. Postle, and J. A. Whitsett. Conditional deletion of abca3 in alveolar type II cells alters surfactant homeostasis in newborn and adult mice. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 298(5):L646–L659, 2010.

[48] M. C. Bewley. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. *Science*, 286(5444):1579–1583, 1999.

[49] H. Bian, Z.-H. Zheng, D. Wei, Z. Zhang, W.-Z. Kang, C.-Q. Hao, K. Dong, W. Kang, J.-L. Xia, J.-L. Miao, R.-H. Xie, B. Wang, X.-X. Sun, X.-M. Yang, P. Lin, J.-J. Geng, K. Wang, H.-Y. Cui, K. Zhang, X.-C. Chen, H. Tang, H. Du, N. Yao, S.-S. Liu, L.-N. Liu, Z. Zhang, Z.-W. Gao, G. Nan, Q.-Y. Wang, J.-Q. Lian, Z.-N. Chen, and P. Zhu. *Meplazumab treats COVID-19 pneumonia: an open-labelled, concurrent controlled add-on clinical trial*. Cold Spring Harbor Laboratory, 2020.

[50] H. M. Biggs, X. Lu, L. Dettinger, S. Sakthivel, J. T. Watson, and S. W. Boktor. Adenovirus-associated influenza-like illness among college students, pennsylvania, USA. *Emerging Infectious Diseases*, 24(11):2117–2119, 2018.

[51] E. Billatos, A. Faiz, Y. Gesthalter, A. LeClerc, Y. O. Alekseyev, X. Xiao, G. Liu, N. H. T. ten Hacken, I. H. Heijink, W. Timens, C. A. Brandsma, D. S. Postma, M. van den Berge, A. Spira, and M. E. Lenburg. Impact of acute exposure to cigarette smoke on airway gene expression. *Physiological Genomics*, 50(9):705–713, 2018.

[52] D. Blanco-Melo, B. E. Nilsson-Payant, W.-C. Liu, S. Uhl, D. Hoagland, R. Møller, T. X. Jordan, K. Oishi, M. Panis, D. Sachs, T. T. Wang, R. E. Schwartz, J. K. Lim, R. A. Albrecht, and B. R. tenOever. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell*, 181(5):1036–1045.e9, 2020.

[53] J. L. Blankman, G. M. Simon, and B. F. Cravatt. A comprehensive profile of brain enzymes that hydrolyze the endocannabinoid 2-arachidonoylglycerol. *Chemistry & Biology*, 14(12):1347–1356, 2007.

[54] K. C. Boardman, A. M. Aryal, W. M. Miller, and C. M. Waters. Actin re-distribution in response to hydrogen peroxide in airway epithelial cells. *Journal of Cellular Physiology*, 199(1):57–66, 2004.

[55] Y. A. Bochkov and J. E. Gern. Rhinoviruses and their receptors: implications for allergic disease. *Current allergy and asthma reports*, 16(4):1–11, 2016.

[56] F. L. Boeuf, J.-S. Diallo, J. A. McCart, S. Thorne, T. Falls, M. Stanford, F. Kanji, R. Auer, C. W. Brown, B. D. Lichty, K. Parato, H. Atkins, D. Kirn, and J. C. Bell. Synergistic interaction between oncolytic viruses augments tumor killing. *Molecular Therapy*, 18(5):888–895, 2010.

[57] D. L. Boggs, J. D. Nguyen, D. Morgenson, M. A. Taffe, and M. Ranganathan. Clinical and preclinical evidence for functional interactions of cannabidiol and δ9-tetrahydrocannabinol. *Neuropsychopharmacology*, 43(1):142–154, 2017.

[58] A. Bondar and J. Lazar. The g protein gi1 exhibits basal coupling but not preassembly with g protein-coupled receptors. *Journal of Biological Chemistry*, 292(23):9690–9698, 2017.

[59] B. Bosier, D. M. Lambert, and E. Hermans. Reciprocal influences of CB1cannabinoid receptor agonists on ERK and JNK signalling in n1e-115 cells. *FEBS Letters*, 582(28):3861–3867, 2008.

[60] R. C. Boswell-Casteel, Y. Fukuda, and J. D. Schuetz. ABCB6, an ABC transporter impacting drug response and disease. *The AAPS Journal*, 20(1), 2017.

[61] I. M. Boudewijn, A. Faiz, K. Steiling, E. van der Wiel, E. Telenga, S. Hoonhorst, N. ten Hacken, C.-A. Brandsma, H. Kerstjens, W. Timens, I. Heijink, M. Jonker, H. de Bruin, S. Vroegop, H. Pasma, W. Boersma, P. Wielders, F. van den Elshout, K. Mansour, A. Spira, M. Lenburg, D. Postma, and M. van den Berge. A nasal gene expression profile differentiates individuals with and without COPD and overlaps bronchial gene expression. In *Genes and Environment*. European Respiratory Society, 2017.

[62] L. J. Buro-Auriemma, J. Salit, N. R. Hackett, M. S. Walters, Y. Strulovici-Barel, M. R. Staudt, J. Fuller, M. Mahmoud, C. S. Stevenson, H. Hilton, M. W. Ho, and R. G. Crystal.

Cigarette smoking induces small airway epithelial epigenetic changes with corresponding modulation of gene expression. *Human Molecular Genetics*, 22(23):4726–4738, 2013.

[63] L. M. Burrell. Myocardial infarction increases ACE2 expression in rat and humans: reply. *European Heart Journal*, 26(11):1142–1143, 2005.

[64] A. Butler, P. Hoffman, P. Smibert, E. Papalexi, and R. Satija. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, 36(5):411–420, 2018.

[65] M. W. Butler, T. Fukui, J. Salit, R. Shaykhiev, J. G. Mezey, N. R. Hackett, and R. G. Crystal. Modulation of cystatin a expression in human airway epithelium related to genotype, smoking, COPD, and lung cancer. *Cancer Research*, 71(7):2572–2581, 2011.

[66] A. Calignano, I. Kátona, F. Désarnaud, A. Giuffrida, G. L. Rana, K. Mackie, T. F. Freund, and D. Piomelli. Bidirectional control of airway responsiveness by endogenous cannabinoids. *Nature*, 408(6808):96–101, 2000.

[67] B. Campling, S. Wright, S. Cole, R. Deeley, and A. Boag. 561 immunohistochemical detection of the multidrug resistance protein (MRP) in human lung cancer. *Lung Cancer*, 18:144, 1997.

[68] H. Canada. Canadian cannabis survey 2017, 2017.

[69] H. Canada. Canadian cannabis survey 2018, 2018.

[70] H. Canada. Canadian cannabis survey 2018, 2019.

[71] Q. T. Cao, J. A. Aguiar, B. J. Tremblay, N. Abbas, N. Tiessen, S. Revill, N. Makhdami, A. Ayoub, G. Cox, K. Ask, et al. Abcf1 regulates dsdna-induced immune responses in human airway epithelial cells. *Frontiers in Cellular and Infection Microbiology*, 10, 2020.

[72] C. R. Carlin. New insights to adenovirus-directed innate immunity in respiratory epithelial cells. *Microorganisms*, 7(8):216, 2019.

[73] B. J. Carolan, B.-G. Harvey, B. P. De, H. Vanni, and R. G. Crystal. Decreased expression of intelectin 1 in the human airway epithelium of smokers compared to nonsmokers. *The Journal of Immunology*, 181(8):5760–5767, 2008.

[74] B. J. Carolan, B.-G. Harvey, N. R. Hackett, T. P. O'Connor, P. A. Cassano, and R. G. Crystal. Disparate oxidant gene expression of airway epithelium compared to alveolar macrophages in smokers. *Respiratory Research*, 10(1), 2009.

[75] B. J. Carolan, A. Heguy, B.-G. Harvey, P. L. Leopold, B. Ferris, and R. G. Crystal. Up-regulation of expression of the ubiquitin carboxyl-terminal hydrolase l1 gene in human airway epithelium of cigarette smokers. *Cancer Research*, 66(22):10729–10740, 2006.

[76] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Di Napoli. Features, evaluation, and treatment of coronavirus (covid-19). *Statpearls [internet]*, 2021.

[77] P. E. Castillo, T. J. Younts, A. E. Chávez, and Y. Hashimotodani. Endocannabinoid signaling and synaptic function. *Neuron*, 76(1):70–81, 2012.

[78] A. B. Chai, A. J. Ammit, and I. C. Gelissen. Examining the role of ABC lipid transporters in pulmonary lipid homeostasis and inflammation. *Respiratory Research*, 18(1), 2017.

[79] T. K. Chan, W. S. D. Tan, H. Y. Peh, and W. S. F. Wong. Aeroallergens induce reactive oxygen species production and DNA damage and dampen antioxidant responses in bronchial epithelial cells. *The Journal of Immunology*, 199(1):39–47, 2017.

[80] H. Chavan, M. Oruganti, and P. Krishnamurthy. The ATP-binding cassette transporter ABCB6 is induced by arsenic and protects against arsenic cytotoxicity. *Toxicological Sciences*, 120(2):519–528, 2011.

[81] S. Cheepala, J.-S. Hulot, J. A. Morgan, Y. Sassi, W. Zhang, A. P. Naren, and J. D. Schuetz. Cyclic nucleotide compartmentalization: Contributions of phosphodiesterases and ATP-binding cassette transporters. *Annual Review of Pharmacology and Toxicology*, 53(1):231–253, 2013.

[82] E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. Meirelles, N. R. Clark, and A. Ma'ayan. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14(1):128, 2013.

[83] W. Chen, H. Srinath, S. S. Lam, C. A. Schiffer, W. E. Royer, and K. Lin. Contribution of ser386 and ser396 to activation of interferon regulatory factor 3. *Journal of Molecular Biology*, 379(2):251–260, 2008.

[84] X. Chen, X. Yang, Y. Zheng, Y. Yang, Y. Xing, and Z. Chen. SARS coronavirus papain-like protease inhibits the type i interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein & Cell*, 5(5):369–381, 2014.

[85] Y.-W. Chen, M.-S. Lee, A. Lucht, F.-P. Chou, W. Huang, T. C. Havighurst, K. Kim, J.-K. Wang, T. M. Antalis, M. D. Johnson, and C.-Y. Lin. TMPRSS2, a serine protease

expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostasomes, is misregulated in prostate cancer cells. *The American Journal of Pathology*, 176(6):2986–2996, 2010.

[86] Z. Chen, L. Mi, J. Xu, J. Yu, X. Wang, J. Jiang, J. Xing, P. Shang, A. Qian, Y. Li, P. X. Shaw, J. Wang, S. Duan, J. Ding, C. Fan, Y. Zhang, Y. Yang, X. Yu, Q. Feng, B. Li, X. Yao, Z. Zhang, L. Li, X. Xue, and P. Zhu. Function of HAb18g/CD147 in invasion of host cells by severe acute respiratory syndrome coronavirus. *The Journal of Infectious Diseases*, 191(5):755–760, 2005.

[87] S. A. Christenson, K. Steiling, M. van den Berge, K. Hijazi, P. S. Hiemstra, D. S. Postma, M. E. Lenburg, A. Spira, and P. G. Woodruff. Asthma–COPD overlap. clinical relevance of genomic signatures of type 2 inflammation in chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 191(7):758–766, 2015.

[88] H. Chu, C.-M. Chan, X. Zhang, Y. Wang, S. Yuan, J. Zhou, R. K.-H. Au-Yeung, K.-H. Sze, D. Yang, H. Shuai, Y. Hou, C. Li, X. Zhao, V. K.-M. Poon, S. P. Leung, M.-L. Yeung, J. Yan, G. Lu, D.-Y. Jin, G. F. Gao, J. F.-W. Chan, and K.-Y. Yuen. Middle east respiratory syndrome coronavirus and bat coronavirus HKU9 both can utilize GRP78 for attachment onto host cells. *Journal of Biological Chemistry*, 293(30):11709–11726, 2018.

[89] R. Cinar, M. R. Iyer, Z. Liu, Z. Cao, T. Jourdan, K. Erdelyi, G. Godlewski, G. Szanda, J. Liu, J. K. Park, B. Mukhopadhyay, A. Z. Rosenberg, J.-S. Lieow, R. G. Lorenz, P. Pacher, R. B. Innis, and G. Kunos. Hybrid inhibitor of peripheral cannabinoid-1 receptors and inducible nitric oxide synthase mitigates liver fibrosis. *JCI Insight*, 1(11), 2016.

[90] G. E. Conner, P. Ivonnet, M. Gelin, P. Whitney, and M. Salathe. H2o2stimulates cystic fibrosis transmembrane conductance regulator through an autocrine prostaglandin pathway, using multidrug-resistant protein–4. *American Journal of Respiratory Cell and Molecular Biology*, 49(4):672–679, 2013.

[91] F. Consortium et al. A promoter-level mammalian expression atlas. *Nature*, 507(7493):462, 2014.

[92] G. O. Consortium. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1):D330–D338, 2018.

[93] A. Corrado, M. Battle, S. K. Wise, F. E.-H. Lee, D. M. Guidot, J. M. DelGaudio, S. A. Molina, and J. M. Levy. Endocannabinoid receptor CB2r is significantly expressed in

aspirin-exacerbated respiratory disease: a pilot study. *International Forum of Allergy & Rhinology*, 8(10):1184–1189, 2018.

[94] K. N. Couper, D. G. Blount, and E. M. Riley. IL-10: The master regulator of immunity to infection. *The Journal of Immunology*, 180(9):5771–5777, 2008.

[95] E. D. Crane, A. A. Al-Hashimi, J. Chen, E. G. Lynn, K. D. Won, Š. Lhoták, M. Naeim, K. Platko, P. Lebeau, J. H. Byun, B. Shayegan, J. C. Krepinsky, K. J. Rayner, S. Marchiò, R. Pasqualini, W. Arap, and R. C. Austin. Anti-GRP78 autoantibodies induce endothelial cell activation and accelerate the development of atherosclerotic lesions. *JCI Insight*, 3(24), 2018.

[96] C. Dai, X. Yao, B. Vaisman, T. Brenner, K. S. Meyer, M. Gao, K. J. Keeran, G. Z. Nugent, X. Qu, Z.-X. Yu, P. K. Dagur, J. P. McCoy, A. T. Remaley, and S. J. Levine. ATP-binding cassette transporter 1 attenuates ovalbumin-induced neutrophilic airway inflammation. *American Journal of Respiratory Cell and Molecular Biology*, 51(5):626–636, 2014.

[97] M. Dai. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Research*, 33(20):e175–e175, 2005.

[98] S. B. Darshini, A. P. Dharshini, and G. R. Kumar. NGS meta data analysis for identification of SNP and INDEL patterns in human airway transcriptome: A preliminary indicator for lung cancer. *Applied & Translational Genomics*, 4:4–9, mar 2015.

[99] S. Davis and P. S. Meltzer. GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. *Bioinformatics*, 23(14):1846–1847, 2007.

[100] R. D. de Vries, A. F. Altenburg, N. J. Nieuwkoop, E. de Bruin, S. E. van Trierum, M. R. Pronk, M. M. Lamers, M. Richard, D. F. Nieuwenhuijse, M. P. G. Koopmans, J. H. C. M. Kreijtz, R. A. M. Fouchier, A. D. M. E. Osterhaus, G. Sutter, and G. F. Rimmelzwaan. Induction of cross-clade antibody and t-cell responses by a modified vaccinia virus ankara-based influenza a(h5n1) vaccine in a randomized phase 1/2a clinical trial. *The Journal of Infectious Diseases*, 218(4):614–623, 2018.

[101] E. de Wit, N. van Doremalen, D. Falzarano, and V. J. Munster. SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology*, 14(8):523–534, 2016.

[102] W. Devane, L. Hanus, A. Breuer, R. Pertwee, L. Stevenson, G. Griffin, D. Gibson, A. Mandelbaum, A. Etinger, and R. Mechoulam. Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science*, 258(5090):1946–1949, 1992.

[103] Z. Ding, L. Fang, H. Jing, S. Zeng, D. Wang, L. Liu, H. Zhang, R. Luo, H. Chen, and S. Xiao. Porcine epidemic diarrhea virus nucleocapsid protein antagonizes beta interferon production by sequestering the interaction between IRF3 and TBK1. *Journal of Virology*, 88(16):8936–8945, 2014.

[104] T. P. Dinh, D. Carpenter, F. M. Leslie, T. F. Freund, I. Katona, S. L. Sensi, S. Kathuria, and D. Piomelli. Brain monoglyceride lipase participating in endocannabinoid inactivation. *Proceedings of the National Academy of Sciences*, 99(16):10819–10824, 2002.

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

[106] M. Donoghue, F. Hsieh, E. Baronas, K. Godbout, M. Gosselin, N. Stagliano, M. Donovan, B. Woolf, K. Robison, R. Jeyaseelan, R. E. Breitbart, and S. Acton. A novel angiotensin-converting enzyme–related carboxypeptidase (ACE2) converts angiotensin i to angiotensin 1-9. *Circulation Research*, 87(5), 2000.

[107] D. Dou, R. Revol, H. Östbye, H. Wang, and R. Daniels. Influenza a virus cell entry, replication, virion assembly and movement. *Frontiers in immunology*, 9:1581, 2018.

[108] S. Dwyer, H. Williams, I. Jones, L. Jones, J. Walters, N. Craddock, M. J. Owen, and M. C. O'Donovan. Investigation of rare non-synonymous variants at ABCA13 in schizophrenia and bipolar disorder. *Molecular Psychiatry*, 16(8):790–791, 2011.

[109] G. D'Amato, L. Cecchi, M. D'Amato, and I. Annesi-Maesano. Climate change and respiratory diseases, 2014.

[110] R. Eccles. Understanding the symptoms of the common cold and influenza. *The Lancet infectious diseases*, 5(11):718–725, 2005.

[111] J. Eddleston, R. U. Lee, A. M. Doerner, J. Herschbach, and B. L. Zuraw. Cigarette smoke decreases innate responses of epithelial cells to rhinovirus infection. *American Journal of Respiratory Cell and Molecular Biology*, 44(1):118–126, 2011.

[112] K. H. Emami, C. Nguyen, H. Ma, D. H. Kim, K. W. Jeong, M. Eguchi, R. T. Moon, J.-L. Teo, S. W. Oh, H. Y. Kim, S. H. Moon, J. R. Ha, and M. Kahn. A small molecule inhibitor of -catenin/CREB-binding protein transcription. *Proceedings of the National Academy of Sciences*, 101(34):12682–12687, 2004.

[113] M. F. Fantauzzi, J. A. Aguiar, B. J.-M. Tremblay, M. J. Mansfield, T. Yanagihara, A. Chandiramohan, S. Revill, M. H. Ryu, C. Carlsten, K. Ask, et al. Expression of endocannabinoid system components in human airway epithelial cells: impact of sex and chronic respiratory disease status. *ERJ open research*, 6(4), 2020.

[114] A. R. Fehr and S. Perlman. Coronaviruses: An overview of their replication and pathogenesis. In *Coronaviruses*, pages 1–23. Springer New York, 2015.

[115] W. R. Fields, J. G. Desiderio, R. M. Leonard, E. E. Burger, B. G. Brown, and D. J. Doolittle. Differential c-myc expression profiles in normal human bronchial epithelial cells following treatment with benzo[a]pyrene, benzo[a]pyrene-4,5 epoxide, and benzo[a]pyrene-7,8-9,10 diol epoxide. *Molecular Carcinogenesis*, 40(2):79–89, 2004.

[116] T. P. Flagg, D. Enkvetchakul, J. C. Koster, and C. G. Nichols. Muscle KATP channels: Recent insights to energy sensing and myoprotection. *Physiological Reviews*, 90(3):799–829, 2010.

[117] S. E. Fligiel, M. D. Roth, E. C. Kleerup, S. H. Barsky, M. S. Simmons, and D. P. Tashkin. Tracheobronchial histopathology in habitual smokers of cocaine, marijuana, and/or tobacco. *Chest*, 112(2):319–326, 1997.

[118] N. C. for Immunization and R. Diseases. Sars basics fact sheet, 2020.

[119] M. W. Foster, W. M. Gwinn, F. L. Kelly, D. M. Brass, A. M. Valente, M. A. Moseley, J. W. Thompson, D. L. Morgan, and S. M. Palmer. Proteomic analysis of primary human airway epithelial cells exposed to the respiratory toxicant diacetyl. *Journal of Proteome Research*, 16(2):538–549, 2017.

[120] Y. Fukuda and J. D. Schuetz. ABC transporters and their role in nucleoside and nucleotide drug resistance. *Biochemical Pharmacology*, 83(8):1073–1083, 2012.

[121] S. Galiegue, S. Mary, J. Marchand, D. Dussossoy, D. Carriere, P. Carayon, M. Bouaboula, D. Shire, G. Fur, and P. Casellas. Expression of central and peripheral cannabinoid receptors in human immune tissues and leukocyte subpopulations. *European Journal of Biochemistry*, 232(1):54–61, 1995.

[122] I. Galve-Roperh, D. Rueda, T. G. del Pulgar, G. Velasco, and M. Guzmán. Mechanism of extracellular signal-regulated kinase activation by the CB1 cannabinoid receptor. *Molecular Pharmacology*, 62(6):1385–1392, 2002.

[123] C. Gao, N. L. Tignor, J. Salit, Y. Strulovici-Barel, N. R. Hackett, R. G. Crystal, and J. G. Mezey. HEFT: eQTL analysis of many thousands of expressed genes while simultaneously controlling for hidden factors. *Bioinformatics*, 30(3):369–376, 2013.

[124] J. Gauthier, A. T. Vincent, S. J. Charette, and N. Derome. A brief history of bioinformatics. *Briefings in bioinformatics*, 20(6):1981–1996, 2019.

[125] P. Gellert, , C. V. Segal, Q. Gao, E. López-Knowles, L.-A. Martin, A. Dodson, T. Li, C. A. Miller, C. Lu, E. R. Mardis, A. Gillman, J. Morden, M. Graf, K. Sidhu, A. Evans, M. Shere, C. Holcombe, S. A. McIntosh, N. Bundred, A. Skene, W. Maxwell, J. Robertson, J. M. Bliss, I. Smith, and M. Dowsett. Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. *Nature Communications*, 7(1), 2016.

[126] M. A. Giembycz, M. Kaur, R. Leigh, and R. Newton. A holy grail of asthma management: toward understanding how long-acting β2-adrenoceptor agonists enhance the clinical efficacy of inhaled corticosteroids. *British journal of pharmacology*, 153(6):1090–1104, 2008.

[127] M. A. Giembycz and R. Newton. Potential mechanisms to explain how labas and pde4 inhibitors enhance the clinical efficacy of glucocorticoids in inflammatory lung diseases. *F1000prime reports*, 7, 2015.

[128] L. Giovannini-Chami, B. Marcet, C. Moreilhon, B. Chevalier, M. I. Illie, K. Lebrigand, K. Robbe-Sermesant, T. Bourrier, J.-F. Michiels, B. Mari, D. Crénesse, P. Hofman, J. de Blic, L. Castillo, M. Albertini, and P. Barbry. Distinct epithelial gene expression phenotypes in childhood respiratory allergy. *European Respiratory Journal*, 39(5):1197–1205, 2011.

[129] A. Girault and E. Brochiero. Evidence of K+ channel function in epithelial cell migration, proliferation, and repair. *American Journal of Physiology-Cell Physiology*, 306(4):C307–C319, 2014.

[130] C. G. Global. regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the global burden of disease study 2017. *Lancet*, 392(10159):1736–1788, 2018.

[131] I. Glowacka, S. Bertram, M. A. Muller, P. Allen, E. Soilleux, S. Pfefferle, I. Steffen, T. S. Tsegaye, Y. He, K. Gnirss, D. Niemeyer, H. Schneider, C. Drosten, and S. Pohlmann. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus

spike protein for membrane fusion and reduces viral control by the humoral immune response. *Journal of Virology*, 85(9):4122–4134, 2011.

[132] M. J. Gold, P. R. Hiebert, H. Y. Park, D. Stefanowicz, A. Le, M. R. Starkey, A. Deane, A. C. Brown, G. Liu, J. C. Horvat, Z. A. Ibrahim, M. B. Sukkar, P. M. Hansbro, C. Carlsten, S. VanEeden, D. D. Sin, K. M. McNagny, D. A. Knight, and J. A. Hirota. Mucosal production of uric acid by airway epithelial cells contributes to particulate matter-induced allergic sensitization. *Mucosal Immunology*, 9(3):809–820, 2015.

[133] G. Gómez-López, J. Dopazo, J. C. Cigudosa, A. Valencia, and F. Al-Shahrour. Precision medicine needs pioneering clinical bioinformaticians. *Briefings in bioinformatics*, 20(3):752–766, 2019.

[134] H. Gong, S. Fligiel, D. P. Tashkin, and R. G. Barbers. Tracheobronchial changes in habitual, heavy smokers of marijuana with and without tobacco. *American Review of Respiratory Disease*, 136(1):142–149, 1987.

[135] M. Gonzalez–Gronow, M. A. Selim, J. Papalas, and S. V. Pizzo. GRP78: A multifunctional receptor on the cell surface. *Antioxidants & Redox Signaling*, 11(9):2299–2306, 2009.

[136] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney, et al. A sars-cov-2 protein interaction map reveals targets for drug repurposing. *Nature*, 583(7816):459–468, 2020.

[137] S. B. Gordon, N. G. Bruce, J. Grigg, P. L. Hibberd, O. P. Kurmi, K. bong Hubert Lam, K. Mortimer, K. P. Asante, K. Balakrishnan, J. Balmes, N. Bar-Zeev, M. N. Bates, P. N. Breysse, S. Buist, Z. Chen, D. Havens, D. Jack, S. Jindal, H. Kan, S. Mehta, P. Moschovis, L. Naeher, A. Patel, R. Perez-Padilla, D. Pope, J. Rylance, S. Semple, and W. J. Martin. Respiratory risks from household air pollution in low and middle income countries. *The Lancet Respiratory Medicine*, 2(10):823–860, 2014.

[138] D. Goubau, S. Deddouche, and C. R. e Sousa. Cytosolic sensing of viruses. *Immunity*, 38(5):855–869, 2013.

[139] A. Granados, A. Peci, A. McGeer, and J. B. Gubbay. Influenza and rhinovirus viral load and disease severity in upper respiratory tract infections. *Journal of Clinical Virology*, 86:14–19, 2017.

[140] S. Grassin-Delyle, E. Naline, A. Buenestado, C. Faisy, J.-C. Alvarez, H. Salvator, C. Abrial, C. Advenier, L. Zemoura, and P. Devillier. Cannabinoids inhibit cholinergic

contraction in human airways through prejunctional CB1receptors. *British Journal of Pharmacology*, 171(11):2767–2777, 2014.

[141] F. Grimminger and R. T. Schermuly. PDGF receptor and its antagonists: Role in treatment of PAH. In *Advances in Experimental Medicine and Biology*, pages 435–446. Humana Press, 2009.

[142] J. Gu and C. Korteweg. Pathology and pathogenesis of severe acute respiratory syndrome. *The American Journal of Pathology*, 170(4):1136–1147, 2007.

[143] Z. Gu, R. Eils, and M. Schlesner. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18):2847–2849, 2016.

[144] A. Gürtler, N. Kunz, M. Gomolka, S. Hornhardt, A. A. Friedl, K. McDonald, J. E. Kohn, and A. Posch. Stain-free technology as a normalization tool in western blot analysis. *Analytical Biochemistry*, 433(2):105–111, 2013.

[145] A. C. Habermann, A. J. Gutierrez, L. T. Bui, S. L. Yahn, N. I. Winters, C. L. Calvi, L. Peter, M.-I. Chung, C. J. Taylor, C. Jetter, L. Raju, J. Roberson, G. Ding, L. Wood, J. M. S. Sucre, B. W. Richmond, A. P. Serezani, W. J. McDonnell, S. B. Mallal, M. J. Bacchetta, J. E. Loyd, C. M. Shaver, L. B. Ware, R. Bremner, R. Walia, T. S. Blackwell, N. E. Banovich, and J. A. Kropski. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science Advances*, 6(28):eaba1972, 2020.

[146] J. Hadjadj, N. Yatim, L. Barnabei, A. Corneau, J. Boussier, N. Smith, H. Péré, B. Charbit, V. Bondet, C. Chenevier-Gobeaux, P. Breillat, N. Carlier, R. Gauzit, C. Morbieu, F. Pène, N. Marin, N. Roche, T.-A. Szwebel, S. H. Merkling, J.-M. Treluyer, D. Veyer, L. Mouthon, C. Blanc, P.-L. Tharaux, F. Rozenberg, A. Fischer, D. Duffy, F. Rieux-Laucat, S. Kernéis, and B. Terrier. Impaired type i interferon activity and inflammatory responses in severe COVID-19 patients. *Science*, 369(6504):718–724, 2020.

[147] J. Hadjadj, N. Yatim, L. Barnabei, A. Corneau, J. Boussier, N. Smith, H. Péré, B. Charbit, V. Bondet, C. Chenevier-Gobeaux, P. Breillat, N. Carlier, R. Gauzit, C. Morbieu, F. Pène, N. Marin, N. Roche, T.-A. Szwebel, S. H. Merkling, J.-M. Treluyer, D. Veyer, L. Mouthon, C. Blanc, P.-L. Tharaux, F. Rozenberg, A. Fischer, D. Duffy, F. Rieux-Laucat, S. Kernéis, and B. Terrier. Impaired type i interferon activity and inflammatory responses in severe COVID-19 patients. *Science*, 369(6504):718–724, 2020.

[148] S. Haga, N. Nagata, T. Okamura, N. Yamamoto, T. Sata, N. Yamamoto, T. Sasazuki, and Y. Ishizaka. TACE antagonists blocking ACE2 shedding caused by the spike protein

of SARS-CoV are candidate antiviral compounds. *Antiviral Research*, 85(3):551–555, 2010.

[149] I. Hamming, W. Timens, M. Bulthuis, A. Lely, G. Navis, and H. van Goor. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. a first step in understanding SARS pathogenesis. *The Journal of Pathology*, 203(2):631–637, 2004.

[150] R. J. Hancox, H. H. Shin, A. R. Gray, R. Poulton, and M. R. Sears. Effects of quitting cannabis on respiratory symptoms. *European Respiratory Journal*, 46(1):80–87, 2015.

[151] B.-G. Harvey, A. Heguy, P. L. Leopold, B. J. Carolan, B. Ferris, and R. G. Crystal. Modification of gene expression of the small airway epithelium in response to cigarette smoking. *Journal of Molecular Medicine*, 85(1):39–53, 2006.

[152] L. E. Haswell, A. Baxter, A. Banerjee, I. Verrastro, J. Mushonganono, J. Adamson, D. Thorne, M. Gaça, and E. Minet. Reduced biological effect of e-cigarette aerosol compared to cigarette smoke evaluated in vitro using normalized nicotine dose and RNA-seq-based toxicogenomics. *Scientific Reports*, 7(1), 2017.

[153] L. E. Haswell, S. Corke, I. Verrastro, A. Baxter, A. Banerjee, J. Adamson, T. Jaunky, C. Proctor, M. Gaça, and E. Minet. In vitro RNA-seq-based toxicogenomics assessment shows reduced biological effect of tobacco heating products when compared to cigarette smoke. *Scientific Reports*, 8(1), 2018.

[154] I. Heijink, A. van Oosterhout, N. Kliphuis, M. Jonker, R. Hoffmann, E. Telenga, K. Klooster, D.-J. Slebos, N. ten Hacken, D. Postma, and M. van den Berge. Oxidant-induced corticosteroid unresponsiveness in human bronchial epithelial cells. *Thorax*, 69(1):5–13, 2013.

[155] C. Henquet, A. Rosa, L. Krabbendam, S. Papiol, L. Faňanás, M. Drukker, J. G. Ramaekers, and J. van Os. An experimental study of catechol-o-methyltransferase val158met moderation of δ-9-tetrahydrocannabinol-induced effects on psychosis and cognition. *Neuropsychopharmacology*, 31(12):2748–2757, 2006.

[156] C. P. Hersh, I. M. Adcock, J. C. Celedón, M. H. Cho, D. C. Christiani, B. E. Himes, N. Kaminski, R. A. Mathias, D. A. Meyers, J. Quackenbush, S. Redline, K. A. Steiling, H. K. Tabor, M. D. Tobin, M. M. Wurfel, I. V. Yang, and G. H. Koppelman. High-throughput sequencing in respiratory, critical care, and sleep medicine research. an official american thoracic society workshop report. *Annals of the American Thoracic Society*, 16(1):1–16, jan 2019.

[157] J. Hessel, J. Heldrich, J. Fuller, M. R. Staudt, S. Radisch, C. Hollmann, B.-G. Harvey, R. J. Kaner, J. Salit, J. Yee-Levin, S. Sridhar, S. Pillai, H. Hilton, G. Wolff, H. Bitter, S. Visvanathan, J. Fine, C. S. Stevenson, R. G. Crystal, and A. E. Tilley. Intraflagellar transport gene expression associated with short cilia in smoking and COPD. *PLoS ONE*, 9(1):e85453, 2014.

[158] A. Heurich, H. Hofmann-Winkler, S. Gierer, T. Liepold, O. Jahn, and S. Pohlmann. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *Journal of Virology*, 88(2):1293–1307, 2013.

[159] P. S. Hiemstra, P. B. McCray, and R. Bals. The innate immune function of airway epithelial cells in inflammatory lung disease. *European Respiratory Journal*, 45(4):1150–1162, 2015.

[160] C. F. Higgins and K. J. Linton. The ATP switch model for ABC transporters. *Nature Structural & Molecular Biology*, 11(10):918–926, 2004.

[161] F. Hikmet, L. Méar, Å. Edvinsson, P. Micke, M. Uhlén, and C. Lindskog. The protein expression profile of ACE2 in human tissues. *Molecular Systems Biology*, 16(7), 2020.

[162] J. A. Hirota, M. J. Gold, P. R. Hiebert, L. G. Parkinson, T. Wee, D. Smith, P. M. Hansbro, C. Carlsten, S. VanEeden, D. D. Sin, K. M. McNagny, and D. A. Knight. The nucleotide-binding domain, leucine-rich repeat protein 3 inflammasome/IL-1 receptor i axis mediates innate, but not adaptive, immune responses after exposure to particulate matter under 10 µm. *American Journal of Respiratory Cell and Molecular Biology*, 52(1):96–105, 2015.

[163] J. A. Hirota, S. A. Hirota, S. M. Warner, D. Stefanowicz, F. Shaheen, P. L. Beck, J. A. MacDonald, T.-L. Hackett, D. D. Sin, S. Van Eeden, et al. The airway epithelium nucleotide-binding domain and leucine-rich repeat protein 3 inflammasome is activated by urban particulate matter. *Journal of Allergy and Clinical Immunology*, 129(4):1116–1125, 2012.

[164] J. A. Hirota and D. A. Knight. Human airway epithelial cell innate immunity: relevance to asthma. *Current Opinion in Immunology*, 24(6):740–746, 2012.

[165] J. A. Hirota, D. J. Marchant, G. K. Singhera, F. Moheimani, D. R. Dorscheid, C. Carlsten, D. Sin, and D. Knight. Urban particulate matter increases human airway epithelial cell IL-1β secretion following scratch wounding and h1n1 influenza a exposure in vitro. *Experimental Lung Research*, 41(6):353–362, 2015.

[166] I. Hlavata, B. Mohelnikova-Duchonova, R. Vaclavikova, V. Liska, P. Pitule, P. Novak, J. Bruha, O. Vycital, L. Holubec, V. Treska, P. Vodicka, and P. Soucek. The role of ABC transporters in progression and clinical outcome of colorectal cancer. *Mutagenesis*, 27(2):187–196, 2012.

[167] D. A. Hoagland, R. Møller, S. A. Uhl, K. Oishi, J. Frere, I. Golynker, S. Horiuchi, M. Panis, D. Blanco-Melo, D. Sachs, et al. Leveraging the antiviral type i interferon system as a first line of defense against sars-cov-2 pathogenicity. *Immunity*, 54(3):557–570, 2021.

[168] R. L. Hodinka. Respiratory RNA viruses. *Microbiology Spectrum*, 4(4), 2016.

[169] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M. A. Müller, C. Drosten, and S. Pöhlmann. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2):271–280.e8, 2020.

[170] J. C. Hogg, P. D. Paré, and T.-L. Hackett. The contribution of small airway obstruction to the pathogenesis of chronic obstructive pulmonary disease. *Physiological Reviews*, 97(2):529–552, 2017.

[171] Y. J. Hou, K. Okuda, C. E. Edwards, D. R. Martinez, T. Asakura, K. H. Dinnon, T. Kato, R. E. Lee, B. L. Yount, T. M. Mascenik, G. Chen, K. N. Olivier, A. Ghio, L. V. Tse, S. R. Leist, L. E. Gralinski, A. Schäfer, H. Dang, R. Gilmore, S. Nakano, L. Sun, M. L. Fulcher, A. Livraghi-Butrico, N. I. Nicely, M. Cameron, C. Cameron, D. J. Kelvin, A. de Silva, D. M. Margolis, A. Markmann, L. Bartelt, R. Zumwalt, F. J. Martinez, S. P. Salvatore, A. Borczuk, P. R. Tata, V. Sontake, A. Kimple, I. Jaspers, W. K. O'Neal, S. H. Randell, R. C. Boucher, and R. S. Baric. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell*, 182(2):429–446.e14, 2020.

[172] S. Hryhorowicz, M. Walczak, O. Zakerska-Banaszak, R. Słomski, and M. Skrzypczak-Zielińska. Pharmacogenetics of cannabinoids. *European Journal of Drug Metabolism and Pharmacokinetics*, 43(1):1–12, 2017.

[173] Y.-H. Hu, Y. Zhang, L.-Q. Jiang, S. Wang, C.-Q. Lei, M.-S. Sun, H.-B. Shu, and Y. Liu. WDFY 1 mediates TLR 3/4 signaling by recruiting TRIF. *EMBO reports*, 16(4):447–455, 2015.

[174] W. Huber, V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K. D. Hansen, R. A. Irizarry,

M. Lawrence, M. I. Love, J. MacDonald, V. Obenchain, A. K. Oleś, H. Pagès, A. Reyes, P. Shannon, G. K. Smyth, D. Tenenbaum, L. Waldron, and M. Morgan. Orchestrating high-throughput genomic analysis with bioconductor. *Nature Methods*, 12(2):115–121, 2015.

[175] M. H. Hudy and D. Proud. Cigarette smoke enhances human rhinovirus-induced CXCL8 production via HuR-mediated mRNA stabilization in human airway epithelial cells. *Respiratory Research*, 14(1):88, 2013.

[176] M. H. Hudy, S. L. Traves, and D. Proud. Transcriptional and epigenetic modulation of human rhinovirus–induced CXCL10 production by cigarette smoke. *American Journal of Respiratory Cell and Molecular Biology*, 50(3):571–582, 2014.

[177] M. H. Hudy, S. L. Traves, S. Wiehler, and D. Proud. Cigarette smoke modulates rhinovirus-induced airway epithelial cell chemokine production. *European Respiratory Journal*, 35(6):1256–1263, 2009.

[178] R. D. Huff, J. A. Aguiar, W. Tse, M. R. Stämpfli, B. J. McConkey, A. C. Doxey, and J. A. Hirota. Effect of long-acting β-agonists/glucocorticoids on human airway epithelial cell cytokine, transcriptomic and oxidative stress responses to cannabis smoke. *ERJ open research*, 6(1), 2020.

[179] R. D. Huff, C. Carlsten, and J. A. Hirota. An update on immunologic mechanisms in the respiratory mucosa in response to air pollutants. *Journal of Allergy and Clinical Immunology*, 143(6):1989–2001, 2019.

[180] R. D. Huff, A. C.-Y. Hsu, K. S. Nichol, B. Jones, D. A. Knight, P. A. B. Wark, P. M. Hansbro, and J. A. Hirota. Regulation of xanthine dehydrogensase gene expression and uric acid production in human airway epithelial cells. *PLOS ONE*, 12(9):e0184260, 2017.

[181] R. D. Huff, C. F. Rider, D. Yan, R. Newton, M. A. Giembycz, C. Carlsten, and J. A. Hirota. Inhibition of ABCC4 potentiates combination beta agonist and glucocorticoid responses in human airway epithelial cells. *Journal of Allergy and Clinical Immunology*, 141(3):1127–1130.e5, 2018.

[182] I. F.-N. Hung, K.-C. Lung, E. Y.-K. Tso, R. Liu, T. W.-H. Chung, M.-Y. Chu, Y.-Y. Ng, J. Lo, J. Chan, A. R. Tam, H.-P. Shum, V. Chan, A. K.-L. Wu, K.-M. Sin, W.-S. Leung, W.-L. Law, D. C. Lung, S. Sin, P. Yeung, C. C.-Y. Yip, R. R. Zhang, A. Y.-F. Fung, E. Y.-W. Yan, K.-H. Leung, J. D. Ip, A. W.-H. Chu, W.-M. Chan, A. C.-K. Ng, R. Lee, K. Fung, A. Yeung, T.-C. Wu, J. W.-M. Chan, W.-W. Yan, W.-M. Chan,

J. F.-W. Chan, A. K.-W. Lie, O. T.-Y. Tsang, V. C.-C. Cheng, T.-L. Que, C.-S. Lau, K.-H. Chan, K. K.-W. To, and K.-Y. Yuen. Triple combination of interferon beta-1b, lopinavir–ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *The Lancet*, 395(10238):1695–1704, 2020.

[183] S. E. Hunt, W. McLaren, L. Gil, A. Thormann, H. Schuilenburg, D. Sheppard, A. Parton, I. M. Armean, S. J. Trevanion, P. Flicek, and F. Cunningham. Ensembl variation resources. *Database*, 2018, 2018.

[184] B. Hurgobin, E. de Jong, and A. Bosco. Insights into respiratory disease through bioinformatics. *Respirology*, 23(12):1117–1126, 2018.

[185] I. M. Ibrahim, D. H. Abdelmalek, M. E. Elshahat, and A. A. Elfiky. COVID-19 spike-host cell receptor GRP78 binding site prediction. *Journal of Infection*, 80(5):554–562, 2020.

[186] H. Ishikawa, Z. Ma, and G. N. Barber. STING regulates intracellular DNA-mediated, type i interferon-dependent innate immunity. *Nature*, 461(7265):788–792, 2009.

[187] M. Iwata, J. E. S. Enciso, and B. H. Greenberg. Selective and specific regulation of ectodomain shedding of angiotensin-converting enzyme 2 by tumor necrosis factor α-converting enzyme. *American Journal of Physiology-Cell Physiology*, 297(5):C1318–C1329, 2009.

[188] K. C. Jamieson, S. L. Traves, C. Kooi, S. Wiehler, C. J. Dumonceaux, B. A. Maciejewski, J. W. Arnason, A. N. Michi, R. Leigh, and D. Proud. Rhinovirus and bacteria synergistically induce IL-17c release from human airway epithelial cells to promote neutrophil recruitment. *The Journal of Immunology*, 202(1):160–170, 2018.

[189] C. A. Janeway and R. Medzhitov. INNATEIMMUNERECOGNITION. *Annual Review of Immunology*, 20(1):197–216, 2002.

[190] Y. M. W. Janssen-Heininger, M. E. Poynter, S. W. Aesif, C. Pantano, J. L. Ather, N. L. Reynaert, K. Ckless, V. Anathy, J. van der Velden, C. G. Irvin, and A. van der Vliet. Nuclear factor b, airway epithelium, and asthma: Avenues for redox control. *Proceedings of the American Thoracic Society*, 6(3):249–255, 2009.

[191] T. Jartti and J. E. Gern. Role of viral infections in the development and exacerbation of asthma in children. *Journal of Allergy and Clinical Immunology*, 140(4):895–906, 2017.

[192] B. Jassal, L. Matthews, G. Viteri, C. Gong, P. Lorente, A. Fabregat, K. Sidiropoulos, J. Cook, M. Gillespie, R. Haw, F. Loney, B. May, M. Milacic, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorser, T. Varusai, J. Weiser, G. Wu, L. Stein, H. Hermjakob, and P. D'Eustachio. The reactome pathway knowledgebase. *Nucleic Acids Research*, 2019.

[193] P. Jha, C. Ramasundarahettige, V. Landsman, B. Rostron, M. Thun, R. N. Anderson, T. McAfee, and R. Peto. 21st-century hazards of smoking and benefits of cessation in the united states. *New England Journal of Medicine*, 368(4):341–350, 2013.

[194] H. P. Jia, D. C. Look, L. Shi, M. Hickey, L. Pewe, J. Netland, M. Farzan, C. Wohlford-Lenane, S. Perlman, and P. B. McCray. ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *Journal of Virology*, 79(23):14614–14621, 2005.

[195] W. E. Johnson, C. Li, and A. Rabinovic. Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics*, 8(1):118–127, 2006.

[196] S. JOUNEAU, N. KHORASANI, P. D. SOUZA, P. MACEDO, J. ZHU, P. K. BHAVSAR, and K. F. CHUNG. EMMPRIN (CD147) regulation of MMP-9 in bronchial epithelial cells in COPD. *Respirology*, 16(4):705–712, 2011.

[197] É. Jubinville, M. Talbot, J.-C. Bérubé, M. Hamel-Auger, M. Maranda-Robitaille, M.-J. Beaulieu, S. Aubin, M.-È. Paré, D. G. Kallend, B. Arsenault, et al. Interplay between cigarette smoking and pulmonary reverse lipid transport. *European Respiratory Journal*, 50(3), 2017.

[198] G. KANE, X. LIU, S. YAMADA, T. OLSON, and A. TERZIC. Cardiac k channels in health and disease. *Journal of Molecular and Cellular Cardiology*, 38(6):937–943, 2005.

[199] M. Kano, T. Ohno-Shosaku, Y. Hashimotodani, M. Uchigashima, and M. Watanabe. Endocannabinoid-mediated control of synaptic transmission. *Physiological Reviews*, 89(1):309–380, 2009.

[200] S. Kathuria, S. Gaetani, D. Fegley, F. Valiño, A. Duranti, A. Tontini, M. Mor, G. Tarzia, G. L. Rana, A. Calignano, A. Giustino, M. Tattoli, M. Palmery, V. Cuomo, and D. Piomelli. Modulation of anxiety through blockade of anandamide hydrolysis. *Nature Medicine*, 9(1):76–81, 2002.

[201] D. Katz, I. Katz, B. Porat-Katz, and Y. Shoenfeld. Medical cannabis: Another piece in the mosaic of autoimmunity? *Clinical Pharmacology & Therapeutics*, 101(2):230–238, 2016.

[202] M. G. Katze, Y. He, and M. Gale. Viruses and interferon: a fight for supremacy. *Nature Reviews Immunology*, 2(9):675–687, 2002.

[203] M. Kaur, J. E. Chivers, M. A. Giembycz, and R. Newton. Long-acting β2-adrenoceptor agonists synergistically enhance glucocorticoid-dependent transcription in human airway epithelial and smooth muscle cells. *Molecular pharmacology*, 73(1):203–214, 2008.

[204] T. Kawai and S. Akira. Innate immune recognition of viral infection. *Nature Immunology*, 7(2):131–137, 2006.

[205] R. KC, S. D. Shukla, S. S. Gautam, P. M. Hansbro, and R. F. O'Toole. The role of environmental exposure to non-cigarette smoke in lung disease. *Clinical and Translational Medicine*, 7(1), 2018.

[206] O. Kebir, G. Lafaye, L. Blecha, B. Chaumette, F. Mouaffak, X. Laqueille, and A. Benyamina. ABCB1 c3435t polymorphism is associated with tetrahydrocannabinol blood levels in heavy cannabis users. *Psychiatry Research*, 262:357–358, 2018.

[207] D. Keppler, I. Leier, G. Jedlitschky, and J. König. ATP-dependent transport of glutathione s-conjugates by the multidrug resistance protein MRP1 and its apical isoform MRP2. *Chemico-Biological Interactions*, 111-112:153–161, 1998.

[208] A. Kicic, E. N. Sutanto, P. T. Stevens, D. A. Knight, and S. M. Stick. Intrinsic biochemical and functional differences in bronchial epithelial cells of children with asthma. *American Journal of Respiratory and Critical Care Medicine*, 174(10):1110–1118, 2006.

[209] J. P. Kiley. Advancing respiratory research. *Chest*, 140(2):497–501, 2011.

[210] H. R. Kim, B. H. Son, S. Y. Lee, K. H. Chung, and S. M. Oh. The role of p53 in marijuana smoke condensates-induced genotoxicity and apoptosis. *Environmental Health and Toxicology*, 27:e2012017, 2012.

[211] M.-S. Kim, S. M. Pinto, D. Getnet, R. S. Nirujogi, S. S. Manda, R. Chaerkady, A. K. Madugundu, D. S. Kelkar, R. Isserlin, S. Jain, et al. A draft map of the human proteome. *Nature*, 509(7502):575–581, 2014.

[212] T. W. Klein. Cannabinoid-based drugs as anti-inflammatory therapeutics. *Nature Reviews Immunology*, 5(5):400–411, 2005.

[213] H. M. Knight, B. S. Pickard, A. Maclean, M. P. Malloy, D. C. Soares, A. F. McRae, A. Condie, A. White, W. Hawkins, K. McGhee, M. van Beck, D. J. MacIntyre, J. M.

Starr, I. J. Deary, P. M. Visscher, D. J. Porteous, R. E. Cannon, D. S. Clair, W. J. Muir, and D. H. Blackwood. A cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. *The American Journal of Human Genetics*, 85(6):833–846, 2009.

[214] N. KOKTURK, G. BOZDAYI, S. YILMAZ, B. DOĞAN, O. GULBAHAR, S. ROTA, and T. TATLICIOGLU. Detection of adenovirus and respiratory syncytial virus in patients with chronic obstructive pulmonary disease: Exacerbation versus stable condition. *Molecular Medicine Reports*, 12(2):3039–3046, 2015.

[215] S. Koyama, K. J. Ishii, C. Coban, and S. Akira. Innate immune response to viral infection. *Cytokine*, 43(3):336–341, 2008.

[216] M. Krassowski, V. Das, S. K. Sahu, and B. B. Misra. State of the field in multi-omics research: From computational needs to data mining and sharing. *Frontiers in Genetics*, 11:1598, 2020.

[217] M. E. Kreft, U. D. Jerman, E. Lasič, N. Hevir-Kene, T. L. Rižner, L. Peternel, and K. Kristan. 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*, 69:1–9, 2015.

[218] R. V. Krieken, N. Mehta, T. Wang, M. Zheng, R. Li, B. Gao, E. Ayaub, K. Ask, J. C. Paton, A. W. Paton, R. C. Austin, and J. C. Krepinsky. Cell surface expression of 78-kDa glucose-regulated protein (GRP78) mediates diabetic nephropathy. *Journal of Biological Chemistry*, 294(19):7755–7768, 2019.

[219] P. C. Krishnamurthy, G. Du, Y. Fukuda, D. Sun, J. Sampath, K. E. Mercer, J. Wang, B. Sosa-Pineda, K. G. Murti, and J. D. Schuetz. Identification of a mammalian mitochondrial porphyrin transporter. *Nature*, 443(7111):586–589, 2006.

[220] F. Krueger. Trim galore!: a wrapper tool around cutadapt and fastqc to consistently apply quality and adapter trimming to fastq files. 0.4. 2015, 2019.

[221] M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, and A. Ma'ayan. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44(W1):W90–W97, 2016.

[222] C.-H. S. Kuo, S. Pavlidis, M. Loza, F. Baribaud, A. Rowe, I. Pandis, U. Hoda, C. Rossios, A. Sousa, S. J. Wilson, et al. A transcriptome-driven analysis of epithelial brushings and bronchial biopsies to define asthma phenotypes in u-biopred. *American journal of respiratory and critical care medicine*, 195(4):443–455, 2017.

[223] O. K. Kurt, J. Zhang, and K. E. Pinkerton. Pulmonary health effects of air pollution. *Current Opinion in Pulmonary Medicine*, 22(2):138–143, 2016.

[224] J. König, A. T. Nies, Y. Cui, I. Leier, and D. Keppler. Conjugate export pumps of the multidrug resistance protein (MRP) family: localization, substrate specificity, and MRP2-mediated drug resistance. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1461(2):377–394, 1999.

[225] M. M. C. Lai. Corona virus: Organization, replication and expression of genome. *Annual Review of Microbiology*, 44(1):303–303, 1990.

[226] A. Larabi, J. M. Devos, S.-L. Ng, M. H. Nanao, A. Round, T. Maniatis, and D. Panne. Crystal structure and mechanism of activation of TANK-binding kinase 1. *Cell Reports*, 3(3):734–746, 2013.

[227] A. J. Lee, G. G. Einarsson, D. F. Gilpin, and M. M. Tunney. Multi-omics approaches: The key to improving respiratory health in people with cystic fibrosis? *Frontiers in pharmacology*, 11:1382, 2020.

[228] M. N. Lee, M. Roy, S.-E. Ong, P. Mertins, A.-C. Villani, W. Li, F. Dotiwala, J. Sen, J. G. Doench, M. H. Orzalli, I. Kramnik, D. M. Knipe, J. Lieberman, S. A. Carr, and N. Hacohen. Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach. *Nature Immunology*, 14(2):179–185, 2012.

[229] J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, 28(6):882–883, 2012.

[230] X. Lei, X. Dong, R. Ma, W. Wang, X. Xiao, Z. Tian, C. Wang, Y. Wang, L. Li, L. Ren, F. Guo, Z. Zhao, Z. Zhou, Z. Xiang, and J. Wang. Activation and evasion of type i interferon responses by SARS-CoV-2. *Nature Communications*, 11(1), 2020.

[231] E. M. Leslie, R. G. Deeley, and S. P. Cole. Multidrug resistance proteins: role of p-glycoprotein, MRP1, MRP2, and BCRP (ABCG2) in tissue defense. *Toxicology and Applied Pharmacology*, 204(3):216–237, 2005.

[232] M. Letko, A. Marzi, and V. Munster. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage b betacoronaviruses. *Nature Microbiology*, 5(4):562–569, 2020.

[233] J. M. Leung, C. X. Yang, A. Tam, T. Shaipanich, T.-L. Hackett, G. K. Singhera, D. R. Dorscheid, and D. D. Sin. ACE-2 expression in the small airway epithelia of smokers and COPD patients: implications for COVID-19. *European Respiratory Journal*, 55(5):2000688, 2020.

[234] C.-X. Li, C. E. Wheelock, C. M. Sköld, and Å. M. Wheelock. Integration of multi-omics datasets enables molecular classification of copd. *European Respiratory Journal*, 51(5), 2018.

[235] H. Li. Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. *arXiv preprint arXiv:1303.3997*, 2013.

[236] H. Li and R. Durbin. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009.

[237] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. D. and. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16):2078–2079, 2009.

[238] W. Li, M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe, and M. Farzan. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, 426(6965):450–454, 2003.

[239] D. Linden, H. Guo-Parke, P. V. Coyle, D. Fairley, D. F. McAuley, C. C. Taggart, and J. Kidney. Respiratory viral infection: a potential "missing" in the pathogenesis of COPD. *European Respiratory Review*, 28(151):180063, 2019.

[240] K. G. Lokugamage, A. Hage, M. de Vries, A. M. Valero-Jimenez, C. Schindewolf, M. Dittmann, R. Rajsbaum, and V. D. Menachery. Type i interferon susceptibility distinguishes sars-cov-2 from sars-cov. *Journal of virology*, 94(23):e01410–20, 2020.

[241] Q.-X. Long, X.-J. Tang, Q.-L. Shi, Q. Li, H.-J. Deng, J. Yuan, J.-L. Hu, W. Xu, Y. Zhang, F.-J. Lv, K. Su, F. Zhang, J. Gong, B. Wu, X.-M. Liu, J.-J. Li, J.-F. Qiu, J. Chen, and A.-L. Huang. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nature Medicine*, 26(8):1200–1204, 2020.

[242] T. W. Loo, M. C. Bartlett, and D. M. Clarke. The "LSGGQ" motif in each nucleotide-binding domain of human p-glycoprotein is adjacent to the opposing walker a sequence. *Journal of Biological Chemistry*, 277(44):41303–41306, 2002.

[243] M. I. Love, W. Huber, and S. Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 2014.

[244] L. Lu, H. Qi, F. Luo, H. Xu, M. Ling, Y. Qin, P. Yang, X. Liu, Q. Yang, J. Xue, C. Chen, J. Lu, Q. Xiang, Q. Liu, and Q. Bian. Feedback circuitry via let-7c between lncRNA CCAT1 and c-myc is involved in cigarette smoke extract-induced malignant transformation of HBE cells. *Oncotarget*, 8(12):19285–19297, 2017.

[245] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W. J. Liu, D. Wang, W. Xu, E. C. Holmes, G. F. Gao, G. Wu, W. Chen, W. Shi, and W. Tan. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224):565–574, 2020.

[246] X. Lu, J. Pan, J. Tao, and D. Guo. SARS-CoV nucleocapsid protein antagonizes IFN-β response by targeting initial step of IFN-β induction pathway, and its c-terminal region is critical for the antagonism. *Virus Genes*, 42(1):37–45, 2010.

[247] C. Lucas, , P. Wong, J. Klein, T. B. R. Castro, J. Silva, M. Sundaram, M. K. Ellingson, T. Mao, J. E. Oh, B. Israelow, T. Takahashi, M. Tokuyama, P. Lu, A. Venkataraman, A. Park, S. Mohanty, H. Wang, A. L. Wyllie, C. B. F. Vogels, R. Earnest, S. Lapidus, I. M. Ott, A. J. Moore, M. C. Muenker, J. B. Fournier, M. Campbell, C. D. Odio, A. Casanovas-Massana, R. Herbst, A. C. Shaw, R. Medzhitov, W. L. Schulz, N. D. Grubaugh, C. D. Cruz, S. Farhadian, A. I. Ko, S. B. Omer, and A. Iwasaki. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature*, 584(7821):463–469, 2020.

[248] P.-Y. Lui, L.-Y. R. Wong, C.-L. Fung, K.-L. Siu, M.-L. Yeung, K.-S. Yuen, C.-P. Chan, P. C.-Y. Woo, K.-Y. Yuen, and D.-Y. Jin. Middle east respiratory syndrome coronavirus m protein suppresses type i interferon expression through the inhibition of TBK1-dependent phosphorylation of IRF3. *Emerging Microbes & Infections*, 5(1):1–9, 2016.

[249] J. Luo, M. Schumacher, A. Scherer, D. Sanoudou, D. Megherbi, T. Davison, T. Shi, W. Tong, L. Shi, H. Hong, C. Zhao, F. Elloumi, W. Shi, R. Thomas, S. Lin, G. Tillinghast,

G. Liu, Y. Zhou, D. Herman, Y. Li, Y. Deng, H. Fang, P. Bushel, M. Woods, and J. Zhang. A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. *The Pharmacogenomics Journal*, 10(4):278–291, 2010.

[250] J. Lynch, Y. Fukuda, P. Krishnamurthy, G. Du, and J. D. Schuetz. Cell survival under stress is enhanced by a mitochondrial ATP-binding cassette transporter that regulates hemoproteins. *Cancer Research*, 69(13):5560–5567, 2009.

[251] M. Mandel, G. Harari, M. Gurevich, and A. Achiron. Cytokine prediction of mortality in COVID19 patients. *Cytokine*, 134:155190, 2020.

[252] W. R. Marrs, J. L. Blankman, E. A. Horne, A. Thomazeau, Y. H. Lin, J. Coy, A. L. Bodor, G. G. Muccioli, S. S.-J. Hu, G. Woodruff, S. Fung, M. Lafourcade, J. P. Alexander, J. Z. Long, W. Li, C. Xu, T. Möller, K. Mackie, O. J. Manzoni, B. F. Cravatt, and N. Stella. The serine hydrolase ABHD6 controls the accumulation and efficacy of 2-AG at cannabinoid receptors. *Nature Neuroscience*, 13(8):951–957, 2010.

[253] C. Mathis, C. Poussin, D. Weisensee, S. Gebel, A. Hengstermann, A. Sewer, V. Belcastro, Y. Xiang, S. Ansari, S. Wagner, J. Hoeng, and M. C. Peitsch. Human bronchial epithelial cells exposed in vitro to cigarette smoke at the air-liquid interface resemble bronchial epithelium from human smokers. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 304(7):L489–L503, 2013.

[254] L. A. Matsuda, S. J. Lolait, M. J. Brownstein, A. C. Young, and T. I. Bonner. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature*, 346(6284):561–564, 1990.

[255] K. Matsumura and S. Ito. Novel biomarker genes which distinguish between smokers and chronic obstructive pulmonary disease patients with machine learning approach. *BMC pulmonary medicine*, 20(1):1–13, 2020.

[256] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, and F. Taguchi. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease tmprss2. *Journal of virology*, 84(24):12658–12664, 2010.

[257] J. E. McDonough, R. Yuan, M. Suzuki, N. Seyednejad, W. M. Elliott, P. G. Sanchez, A. C. Wright, W. B. Gefter, L. Litzky, H. O. Coxson, P. D. Paré, D. D. Sin, R. A. Pierce, J. C. Woods, A. M. McWilliams, J. R. Mayo, S. C. Lam, J. D. Cooper, and J. C. Hogg. Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. *New England Journal of Medicine*, 365(17):1567–1575, 2011.

[258] L. P. McGarvey, C. A. Butler, S. Stokesberry, L. Polley, S. McQuaid, H. Abdullah, S. Ashraf, M. K. McGahon, T. M. Curtis, J. Arron, D. Choy, T. J. Warke, P. Bradding, M. Ennis, A. Zholos, R. W. Costello, and L. G. Heaney. Increased expression of bronchial epithelial transient receptor potential vanilloid 1 channels in patients with severe asthma. *Journal of Allergy and Clinical Immunology*, 133(3):704–712.e4, 2014.

[259] D. Merico, R. Isserlin, O. Stueker, A. Emili, and G. D. Bader. Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE*, 5(11):e13984, 2010.

[260] E. V. Mesev, R. A. LeDesma, and A. Ploss. Decoding type i and III interferon signalling during viral infection. *Nature Microbiology*, 4(6):914–924, 2019.

[261] R. L. Minaker, K. L. Mossman, and J. R. Smiley. Functional inaccessibility of quiescent herpes simplex virus genomes. *Virology Journal*, 2(1):85, 2005.

[262] L. Miyashita and G. Foley. E-cigarettes and respiratory health: the latest evidence. *The Journal of Physiology*, 598(22):5027–5038, 2020.

[263] B. D. Modena, J. R. Tedrow, J. Milosevic, E. R. Bleecker, D. A. Meyers, W. Wu, Z. Bar-Joseph, S. C. Erzurum, B. M. Gaston, W. W. Busse, N. N. Jarjour, N. Kaminski, and S. E. Wenzel. Gene expression in relation to exhaled nitric oxide identifies novel asthma phenotypes with unique biomolecular pathways. *American Journal of Respiratory and Critical Care Medicine*, 190(12):1363–1372, 2014.

[264] M. A. Modestou, L. J. Manzel, S. El-Mahdy, and D. C. Look. Inhibition of IFN-γ-dependent antiviral airway epithelial defense by cigarette smoke. *Respiratory Research*, 11(1), 2010.

[265] F. Moheimani, H. M. Roth, J. Cross, A. T. Reid, F. Shaheen, S. M. Warner, J. A. Hirota, A. Kicic, T. S. Hallstrand, M. Kahn, S. M. Stick, P. M. Hansbro, T.-L. Hackett, and D. A. Knight. Disruption of β-catenin/CBP signaling inhibits human airway epithelial–mesenchymal transition and repair. *The International Journal of Biochemistry & Cell Biology*, 68:59–69, 2015.

[266] D. Moir, W. S. Rickert, G. Levasseur, Y. Larose, R. Maertens, P. White, and S. Desjardins. A comparison of mainstream and sidestream marijuana and tobacco cigarette smoke produced under two machine smoking conditions. *Chemical Research in Toxicology*, 21(2):494–502, 2008.

[267] B. Moldoveanu, P. Otmishi, P. Jani, J. Walker, X. Sarmiento, J. Guardiola, M. Saad, and J. Yu. Inflammatory mechanisms in the lung. *Journal of inflammation research*, 2:1, 2009.

[268] S. Munro, K. L. Thomas, and M. Abu-Shaar. Molecular characterization of a peripheral receptor for cannabinoids. *Nature*, 365(6441):61–65, 1993.

[269] V. J. Munster, M. Koopmans, N. van Doremalen, D. van Riel, and E. de Wit. A novel coronavirus emerging in china — key questions for impact assessment. *New England Journal of Medicine*, 382(8):692–694, 2020.

[270] N. S. Nagathihalli, P. P. Massion, A. L. Gonzalez, P. Lu, and P. K. Datta. Smoking induces epithelial-to-mesenchymal transition in non–small cell lung cancer through HDAC-mediated downregulation of e-cadherin. *Molecular Cancer Therapeutics*, 11(11):2362–2372, 2012.

[271] J. A. Nasir, R. A. Kozak, P. Aftanas, A. R. Raphenya, K. M. Smith, F. Maguire, H. Maan, M. Alruwaili, A. Banerjee, H. Mbareche, B. P. Alcock, N. C. Knox, K. Mossman, B. Wang, J. A. Hiscox, A. G. McArthur, and S. Mubareka. A comparison of whole genome sequencing of SARS-CoV-2 using amplicon-based sequencing, random hexamers, and bait capture. *Viruses*, 12(8):895, 2020.

[272] R. Nielsen, J. S. Paul, A. Albrechtsen, and Y. S. Song. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12(6):443–451, may 2011.

[273] D. Niemeyer, T. Zillinger, D. Muth, F. Zielecki, G. Horvath, T. Suliman, W. Barchet, F. Weber, C. Drosten, and M. A. Muller. Middle east respiratory syndrome coronavirus accessory protein 4a is a type i interferon antagonist. *Journal of Virology*, 87(22):12489–12495, 2013.

[274] S. Nish and R. Medzhitov. Host defense pathways: role of redundancy and compensation in infectious disease phenotypes. *Immunity*, 34(5):629–636, 2011.

[275] R. S. Noyce, S. E. Collins, and K. L. Mossman. Differential modification of interferon regulatory factor 3 following virus particle entry. *Journal of Virology*, 83(9):4013–4022, 2009.

[276] S. L. O'Beirne, S. A. Shenoy, J. Salit, Y. Strulovici-Barel, R. J. Kaner, S. Visvanathan, J. S. Fine, J. G. Mezey, and R. G. Crystal. Ambient pollution–related reprogramming of the human small airway epithelial transcriptome. *American Journal of Respiratory and Critical Care Medicine*, 198(11):1413–1422, 2018.

[277] U. D. of Health, C. f. D. C. Human Services, N. C. f. C. D. P. Prevention, O. o. S. Health Promotion, and Health. The health consequences of smoking—50 years of progress: A report of the surgeon general. Report, U.S. Department of Health and Human Services, 2014.

[278] O. E. Onwuameze, K. W. Nam, E. A. Epping, T. H. Wassink, S. Ziebell, N. C. Andreasen, and B.-C. Ho. MAPK14 and CNR1 gene variant interactions: effects on brain volume deficits in schizophrenia patients with marijuana misuse. *Psychological Medicine*, 43(3):619–631, 2012.

[279] W. H. Organization. World drug report 2018, 2018.

[280] W. H. Organization. Cancer. Report, World Health Organization, 2020.

[281] W. H. Organization. Cannabis. Report, World Health Organization, 2021.

[282] W. H. Organization. Tobacco. Report, World Health Organization, 2021.

[283] P. Pacher and G. Kunos. Modulating the endocannabinoid system in human health and disease - successes and failures. *FEBS Journal*, 280(9):1918–1943, 2013.

[284] M. Paczkowska, , J. Barenboim, N. Sintupisut, N. S. Fox, H. Zhu, D. Abd-Rabbo, M. W. Mee, P. C. Boutros, and J. R. and. Integrative pathway enrichment analysis of multivariate omics data. *Nature Communications*, 11(1), 2020.

[285] U. Pagotto, G. Marsicano, D. Cota, B. Lutz, and R. Pasquali. The emerging role of the endocannabinoid system in endocrine regulation and energy balance. *Endocrine Reviews*, 27(1):73–100, 2005.

[286] R. S. Pappas, M. R. Fresquez, N. Martone, and C. H. Watson. Toxic metal concentrations in mainstream smoke from cigarettes available in the USA. *Journal of Analytical Toxicology*, 38(4):204–211, 2014.

[287] A. Park and A. Iwasaki. Type i and type III interferons – induction, signaling, evasion, and application to combat COVID-19. *Cell Host & Microbe*, 27(6):870–878, 2020.

[288] D. Parker and A. Prince. Innate immunity in the respiratory epithelium. *American Journal of Respiratory Cell and Molecular Biology*, 45(2):189–201, 2011.

[289] M. M. Parker, R. P. Chase, A. Lamb, A. Reyes, A. Saferali, J. H. Yun, B. E. Himes, E. K. Silverman, C. P. Hersh, and P. J. Castaldi. Rna sequencing identifies novel non-coding rna and exon-specific effects associated with cigarette smoking. *BMC medical genomics*, 10(1):1–10, 2017.

[290] J. K. Paterson, S. Shukla, C. M. Black, T. Tachiwada, S. Garfield, S. Wincovitch, D. N. Ernst, A. Agadir, X. Li, S. V. Ambudkar, G. Szakacs, S. ichi Akiyama, and M. M. Gottesman. Human ABCB6 localizes to both the outer mitochondrial membrane and the plasma membrane. *Biochemistry*, 46(33):9443–9452, 2007.

[291] R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4):417–419, 2017.

[292] R. A. Pauwels, C.-G. Löfdahl, D. S. Postma, A. E. Tattersfield, P. O'Byrne, P. J. Barnes, and A. Ullman. Effect of inhaled formoterol and budesonide on exacerbations of asthma. *New England Journal of Medicine*, 337(20):1405–1411, 1997.

[293] R. Pawliczak, C. Logun, P. Madara, J. Barb, A. F. Suffredini, P. J. Munson, R. L. Danner, and J. H. Shelhamer. Influence of IFN-γ on gene expression in normal human bronchial epithelial cells: modulation of IFN-γ effects by dexamethasone. *Physiological Genomics*, 23(1):28–45, 2005.

[294] S. Paytubi, N. A. Morrice, J. Boudeau, and C. G. Proud. The n-terminal region of ABC50 interacts with eukaryotic initiation factor eIF2 and is a target for regulatory phosphorylation by CK2. *Biochemical Journal*, 409(1):223–231, 2007.

[295] S. Paytubi, X. Wang, Y. W. Lam, L. Izquierdo, M. J. Hunter, E. Jan, H. S. Hundal, and C. G. Proud. ABC50 promotes translation initiation in mammalian cells. *Journal of Biological Chemistry*, 284(36):24061–24073, 2009.

[296] J. S. Peiris, K. Y. Yuen, A. D. Osterhaus, and K. Stöhr. The severe acute respiratory syndrome. *New England Journal of Medicine*, 349(25):2431–2441, 2003.

[297] R. Pereda, D. González, H. B. Rivero, J. C. Rivero, A. Pérez, L. D. R. Lopez, N. Mezquia, R. Venegas, J. R. Betancourt, R. E. Domínguez, et al. Therapeutic effectiveness of interferon alpha 2b treatment for covid-19 patient recovery. *Journal of Interferon & Cytokine Research*, 40(12):578–588, 2020.

[298] C. C. Perez. *Invisible Women: Data Bias in A World Designed for Men*. Abrams Press, 2019.

[299] S. Perlman and J. Netland. Coronaviruses post-SARS: update on replication and pathogenesis. *Nature Reviews Microbiology*, 7(6):439–450, 2009.

[300] R. G. Pertwee, A. C. Howlett, M. E. Abood, S. P. H. Alexander, V. D. Marzo, M. R. Elphick, P. J. Greasley, H. S. Hansen, G. Kunos, K. Mackie, R. Mechoulam, and R. A. Ross. International union of basic and clinical pharmacology. LXXIX. cannabinoid receptors and their ligands: Beyond CB1and CB2. *Pharmacological Reviews*, 62(4):588–631, 2010.

[301] S. R. Piccolo, Y. Sun, J. D. Campbell, M. E. Lenburg, A. H. Bild, and W. E. Johnson. A single-sample microarray normalization method to facilitate personalized-medicine workflows. *Genomics*, 100(6):337–344, 2012.

[302] A. Pilz, K. Ramsauer, H. Heidari, M. Leitges, P. Kovarik, and T. Decker. Phosphorylation of the stat1 transactivating domain is required for the response to type i interferons. *EMBO reports*, 4(4):368–373, 2003.

[303] K. E. Pinkerton, L. S. V. Winkle, C. G. Plopper, S. Smiley-Jewell, E. C. Covarrubias, and J. T. McBride. Architecture of the trachobronchial tree. In R. A. Parent, editor, *Comparative Biology of the Normal Lung*, chapter 4, pages 33–51. Elsevier Inc., 2015.

[304] Ż. Piotrowska, M. Niezgoda, W. Łebkowski, A. Filipek, N. Domian, and I. Kasacka. Sex differences in distribution of cannabinoid receptors (CB1 and CB2), s100a6 and CacyBP/SIP in human ageing hearts. *Biology of Sex Differences*, 9(1), 2018.

[305] L. W. Plasschaert, R. Žilionis, R. Choo-Wing, V. Savova, J. Knehr, G. Roma, A. M. Klein, and A. B. Jaffe. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature*, 560(7718):377–381, 2018.

[306] D. S. Postma. Gender differences in asthma development and progression. *Gender Medicine*, 4:S133–S146, 2007.

[307] T. Powles, R. te Poele, J. Shamash, T. Chaplin, D. Propper, S. Joel, T. Oliver, and W. M. Liu. Cannabis-induced cytotoxicity in leukemic cell lines: the role of the cannabinoid receptors and the MAPK pathway. *Blood*, 105(3):1214–1221, 2005.

[308] A. H. Pripp and M. Stanišić. The correlation between pro- and anti-inflammatory cytokines in chronic subdural hematoma patients assessed with factor analysis. *PLoS ONE*, 9(2):e90149, 2014.

[309] R. N. Proctor. The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll. *Tobacco control*, 21(2):87–91, 2012.

[310] D. Proud and R. Leigh. Epithelial cells and airway diseases. *Immunological Reviews*, 242(1):186–204, 2011.

[311] F. Qi, S. Qian, S. Zhang, and Z. Zhang. *Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses*. Cold Spring Harbor Laboratory, 2020.

[312] T. Raman, T. P. O'Connor, N. R. Hackett, W. Wang, B.-G. Harvey, M. A. Attiyeh, D. T. Dang, M. Teater, and R. G. Crystal. Quality control in microarray assessment of gene expression in human airway epithelium. *BMC Genomics*, 10(1):493, 2009.

[313] R. D. Ramirez, S. Sheridan, L. Girard, M. Sato, Y. Kim, J. Pollack, M. Peyton, Y. Zou, J. M. Kurie, J. M. DiMaio, S. Milchgrub, A. L. Smith, R. F. Souza, L. Gilbey, X. Zhang, K. Gandia, M. B. Vaughan, W. E. Wright, A. F. Gazdar, J. W. Shay, and J. D. Minna. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Research*, 64(24):9027–9034, 2004.

[314] A. Rebendenne, A. L. Chaves Valadão, M. Tauziet, G. Maarifi, B. Bonaventure, J. McKellar, R. Planès, S. Nisole, M. Arnaud-Arnould, O. Moncorgé, et al. Sars-cov-2 triggers an mda-5-dependent interferon response which is unable to control replication in lung epithelial cells. *Journal of Virology*, 95(8):e02415–20, 2021.

[315] J. Reimand, R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Rostamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, D. Merico, and G. D. Bader. Pathway enrichment analysis and visualization of omics data using g:profiler, GSEA, cytoscape and EnrichmentMap. *Nature Protocols*, 14(2):482–517, 2019.

[316] S. D. Reynolds, K. E. Pinkerton, and A. R. Mariassy. Epithelial cells of trachea and bronchi. In R. A. Parent, editor, *Comparative Biology of the Normal Lung*, chapter 6, pages 61–81. Elsevier Inc., 2015.

[317] L. I. Ribeiro and P. W. Ind. Effect of cannabis smoking on lung function and respiratory symptoms: a structured literature review. *npj Primary Care Respiratory Medicine*, 26(1), 2016.

[318] D. P. Rich, M. P. Anderson, R. J. Gregory, S. H. Cheng, S. Paul, D. M. Jefferson, J. D. McCann, K. W. Klinger, A. E. Smith, and M. J. Welsh. Expression of cystic fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells. *Nature*, 347(6291):358–363, 1990.

[319] M. Richard, R. Drouin, and A. D. Beaulieu. ABC50, a novel human ATP-binding cassette protein found in tumor necrosis factor-α-stimulated synoviocytes. *Genomics*, 53(2):137–145, 1998.

[320] C. F. Rider, E. M. King, N. S. Holden, M. A. Giembycz, and R. Newton. Inflammatory stimuli inhibit glucocorticoid-dependent transactivation in human pulmonary epithelial cells: Rescue by long-acting β2-adrenoceptor agonists. *Journal of Pharmacology and Experimental Therapeutics*, 338(3):860–869, 2011.

[321] M. A. Riedl and A. E. Nel. Importance of oxidative stress in the pathogenesis and treatment of asthma. *Current Opinion in Allergy & Clinical Immunology*, 8(1):49–56, 2008.

[322] T. N. Rindler, C. A. Stockman, A. L. Filuta, K. M. Brown, J. M. Snowball, W. Zhou, R. Veldhuizen, E. M. Zink, S. E. Dautel, G. Clair, C. Ansong, Y. Xu, J. P. Bridges, and J. A. Whitsett. Alveolar injury and regeneration following deletion of ABCA3. *JCI Insight*, 2(24), 2017.

[323] J. Riordan. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Trends in Genetics*, 5:363, 1989.

[324] A. I. Ritchie, H. A. Farne, A. Singanayagam, D. J. Jackson, P. Mallia, and S. L. Johnston. Pathogenesis of viral infection in exacerbations of airway disease. *Annals of the American Thoracic Society*, 12(Supplement 2):S115–S132, 2015.

[325] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7):e47–e47, 2015.

[326] M. ROTH, A. ARORA, S. BARSKY, E. KLEERUP, M. SIMMONS, and D. TASHKIN. Airway inflammation in young marijuana and tobacco smokers. *American Journal of Respiratory and Critical Care Medicine*, 157(3):928–937, 1998.

[327] M. D. Roth, J. A. Marques-Magallanes, M. Yuan, W. Sun, D. P. Tashkin, and O. Hankinson. Induction and regulation of the carcinogen-metabolizing enzyme cyp1a1 by marijuana smoke and δ9-tetrahydrocannabinol. *American Journal of Respiratory Cell and Molecular Biology*, 24(3):339–344, 2001.

[328] L. R. Ruhaak, J. Felth, P. C. Karlsson, J. J. Rafter, R. Verpoorte, and L. Bohlin. Evaluation of the cyclooxygenase inhibiting effects of six major cannabinoids isolated from cannabis sativa. *Biological and Pharmaceutical Bulletin*, 34(5):774–778, 2011.

[329] E. Ryberg, N. Larsson, S. Sjögren, S. Hjorth, N.-O. Hermansson, J. Leonova, T. Elebring, K. Nilsson, T. Drmota, and P. J. Greasley. The orphan receptor GPR55 is a novel cannabinoid receptor. *British Journal of Pharmacology*, 152(7):1092–1101, 2007.

[330] M. A. Sadiq, M. Hanout, S. Sarwar, M. Hassan, D. V. Do, Q. D. Nguyen, and Y. J. Sepah. Platelet derived growth factor inhibitors: A potential therapeutic approach for ocular neovascularization. *Saudi Journal of Ophthalmology*, 29(4):287–291, 2015.

[331] S. P. Saha, D. K. Bhalla, T. F. Whayne, and C. G. Gairola. Cigarette smoke and adverse health effects: An overview of research trends and future needs. *International Journal of Angiology*, 16(03):77–83, 2007.

[332] F. Sanger, G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. Hutchison, P. M. Slocombe, and M. Smith. Nucleotide sequence of bacteriophage φx174 dna. *nature*, 265(5596):687–695, 1977.

[333] S. G. Sawicki, D. L. Sawicki, and S. G. Siddell. A contemporary view of coronavirus transcription. *Journal of Virology*, 81(1):20–29, 2006.

[334] A. C. Schamberger, N. Mise, J. Jia, E. Genoyer, A. O. Yildirim, S. Meiners, and O. Eickelberg. Cigarette smoke–induced disruption of bronchial epithelial tight junctions is prevented by transforming growth factor-β. *American Journal of Respiratory Cell and Molecular Biology*, 50(6):1040–1052, 2014.

[335] F. Schembri, S. Sridhar, C. Perdomo, A. M. Gustafson, X. Zhang, A. Ergun, J. Lu, G. Liu, X. Zhang, J. Bowers, C. Vaziri, K. Ott, K. Sensinger, J. J. Collins, J. S. Brody, R. Getts, M. E. Lenburg, and A. Spira. MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. *Proceedings of the National Academy of Sciences*, 106(7):2319–2324, 2009.

[336] H. B. Schiller, D. T. Montoro, L. M. Simon, E. L. Rawlins, K. B. Meyer, M. Strunz, F. A. V. Braga, W. Timens, G. H. Koppelman, G. R. S. Budinger, J. K. Burgess, A. Waghray, M. van den Berge, F. J. Theis, A. Regev, N. Kaminski, J. Rajagopal, S. A. Teichmann, A. V. Misharin, and M. C. Nawijn. The human lung cell atlas: A high-resolution reference map of the human lung in health and disease. *American Journal of Respiratory Cell and Molecular Biology*, 61(1):31–41, 2019.

[337] J. W. Schoggins. Interferon-stimulated genes: What do they all do? *Annual Review of Virology*, 6(1):567–584, 2019.

[338] J. W. Schoggins and C. M. Rice. Interferon-stimulated genes and their antiviral effector functions. *Current Opinion in Virology*, 1(6):519–525, 2011.

[339] K. S. Schulz and K. L. Mossman. Viral evasion strategies in type i IFN signaling – a summary of recent developments. *Frontiers in Immunology*, 7, 2016.

[340] J. Severin, , M. Lizio, J. Harshbarger, H. Kawaji, C. O. Daub, Y. Hayashizaki, N. Bertin, and A. R. R. Forrest. Interactive visualization and analysis of large-scale sequencing datasets using ZENBU. *Nature Biotechnology*, 32(3):217–219, 2014.

[341] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, and F. Li. Structural basis of receptor recognition by sars-cov-2. *Nature*, 581(7807):221–224, 2020.

[342] V. Shang, S. O'Sullivan, D. Kendall, and R. Roberts. The endogenous cannabinoid anandamide increases human airway epithelial cell permeability through an arachidonic acid metabolite. *Pharmacological Research*, 105:152–163, 2016.

[343] V. C. Shang, D. A. Kendall, and R. E. Roberts. δ9-tetrahydrocannabinol reverses TNFα-induced increase in airway epithelial cell permeability through CB2 receptors. *Biochemical Pharmacology*, 120:63–71, 2016.

[344] P. Shannon. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498–2504, 2003.

[345] S. Sharma, T. L. Phang, J. Gaydos, A. McNally, and E. L. Burnham. Differential gene expression between cannabis and tobacco smoke exposure. In *C21. OMICS IN LUNG DISEASE*, pages A4966–A4966. American Thoracic Society, 2017.

[346] R. Shaykhiev, F. Otaki, P. Bonsu, D. T. Dang, M. Teater, Y. Strulovici-Barel, J. Salit, B.-G. Harvey, and R. G. Crystal. Cigarette smoking reprograms apical junctional complex molecular architecture in the human airway epithelium in vivo. *Cellular and Molecular Life Sciences*, 68(5):877–892, 2010.

[347] D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, A. Bhattacharya, L. Schulz, M. Widera, A. R. Mehdipour, G. Tascher, P. P. Geurink, A. Wilhelm, G. J. van der Heden van Noort, H. Ovaa, S. Müller, K.-P. Knobeloch, K. Rajalingam, B. A. Schulman, J. Cinatl, G. Hummer, S. Ciesek, and I. Dikic. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*, 587(7835):657–662, 2020.

[348] S. Shulenin, L. M. Nogee, T. Annilo, S. E. Wert, J. A. Whitsett, and M. Dean. ABCA3gene mutations in newborns with fatal surfactant deficiency. *New England Journal of Medicine*, 350(13):1296–1303, 2004.

[349] A. Shulla, T. Heald-Sargent, G. Subramanya, J. Zhao, S. Perlman, and T. Gallagher. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *Journal of Virology*, 85(2):873–882, 2010.

[350] G. Simmons, D. N. Gosalia, A. J. Rennekamp, J. D. Reeves, S. L. Diamond, and P. Bates. Inhibitors of cathepsin l prevent severe acute respiratory syndrome coronavirus entry. *Proceedings of the National Academy of Sciences*, 102(33):11876–11881, 2005.

[351] A. C. Sims, R. S. Baric, B. Yount, S. E. Burkett, P. L. Collins, and R. J. Pickles. Severe acute respiratory syndrome coronavirus infection of human ciliated airway epithelia: Role of ciliated cells in viral spread in the conducting airways of the lungs. *Journal of Virology*, 79(24):15511–15524, 2005.

[352] A. Singanayagam, N. Glanville, J. L. Girkin, Y. M. Ching, A. Marcellini, J. D. Porter, M. Toussaint, R. P. Walton, L. J. Finney, J. Aniscenko, J. Zhu, M.-B. Trujillo-Torralbo, M. A. Calderazzo, C. Grainge, S.-L. Loo, P. C. Veerati, P. S. Pathinayake, K. S. Nichol, A. T. Reid, P. L. James, R. Solari, P. A. B. Wark, D. A. Knight, M. F. Moffatt, W. O. Cookson, M. R. Edwards, P. Mallia, N. W. Bartlett, and S. L. Johnston. Corticosteroid suppression of antiviral immunity increases bacterial loads and mucus production in COPD exacerbations. *Nature Communications*, 9(1), 2018.

[353] K. P. Singh, K. P. Maremanda, D. Li, and I. Rahman. Exosomal micrornas are novel circulating biomarkers in cigarette, waterpipe smokers, e-cigarette users and dual smokers. *BMC medical genomics*, 13(1):1–20, 2020.

[354] S. Singh, S. Vrishni, B. K. Singh, I. Rahman, and P. Kakkar. Nrf2-ARE stress response mechanism: A control point in oxidative stress-mediated dysfunctions and chronic inflammatory diseases. *Free Radical Research*, 44(11):1267–1288, 2010.

[355] K.-L. Siu, M. L. Yeung, K.-H. Kok, K.-S. Yuen, C. Kew, P.-Y. Lui, C.-P. Chan, H. Tse, P. C. Y. Woo, K.-Y. Yuen, and D.-Y. Jin. Middle east respiratory syndrome coronavirus 4a protein is a double-stranded RNA-binding protein that suppresses PACT-induced activation of RIG-i and MDA5 in the innate antiviral response. *Journal of Virology*, 88(9):4866–4876, 2014.

[356] C. Soneson, M. I. Love, and M. D. Robinson. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, 4:1521, 2015.

[357] J. Sonett, M. Goldklang, P. Sklepkiewicz, A. Gerber, J. Trischler, T. Zelonina, M. Westerterp, V. Lemaître, Y. Okada, and J. D. Armiento. A critical role for abc transporters

in persistent lung inflammation in the development of emphysema after smoke exposure. *The FASEB Journal*, 32(12):6724–6736, 2018.

[358] I.-C. Sorheim, A. Johannessen, A. Gulsvik, P. S. Bakke, E. K. Silverman, and D. L. DeMeo. Gender differences in COPD: are women more susceptible to smoking effects than men? *Thorax*, 65(6):480–485, 2010.

[359] A. Spira, J. Beane, V. Shah, G. Liu, F. Schembri, X. Yang, J. Palma, and J. S. Brody. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proceedings of the National Academy of Sciences*, 101(27):10143–10148, 2004.

[360] M. D. Srivastava, B. Srivastava, and B. Brouhard. $\delta 9$ tetrahydrocannabinol and cannabidiol alter cytokine production by human immune cells. *Immunopharmacology*, 40(3):179–185, 1998.

[361] W. Stacklies, H. Redestig, M. Scholz, D. Walther, and J. Selbig. pcaMethods a bioconductor package providing PCA methods for incomplete data. *Bioinformatics*, 23(9):1164–1167, 2007.

[362] M. R. Stämpfli and G. P. Anderson. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. *Nature Reviews Immunology*, 9(5):377–384, 2009.

[363] H. C. Steen and A. M. Gamero. STAT2 phosphorylation and signaling. *JAK-STAT*, 2(4):e25790, 2013.

[364] K. Steiling, M. van den Berge, K. Hijazi, R. Florido, J. Campbell, G. Liu, J. Xiao, X. Zhang, G. Duclos, E. Drizik, H. Si, C. Perdomo, C. Dumont, H. O. Coxson, Y. O. Alekseyev, D. Sin, P. Pare, J. C. Hogg, A. McWilliams, P. S. Hiemstra, P. J. Sterk, W. Timens, J. T. Chang, P. Sebastiani, G. T. O'Connor, A. H. Bild, D. S. Postma, S. Lam, A. Spira, and M. E. Lenburg. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. *American Journal of Respiratory and Critical Care Medicine*, 187(9):933–942, 2013.

[365] N. Stella, P. Schweitzer, and D. Piomelli. A second endogenous cannabinoid that modulates long-term potentiation. *Nature*, 388(6644):773–778, 1997.

[366] A. Stentebjerg-Andersen, I. V. Notlevsen, B. Brodin, and C. U. Nielsen. Calu-3 cells grown under AIC and LCC conditions: Implications for dipeptide uptake and transepithelial transport of substances. *European Journal of Pharmaceutics and Biopharmaceutics*, 78(1):19–26, 2011.

[367] D. B. Stetson and R. Medzhitov. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity*, 24(1):93–103, 2006.

[368] Y. Strulovici-Barel, L. Omberg, M. O'Mahony, C. Gordon, C. Hollmann, A. E. Tilley, J. Salit, J. Mezey, B.-G. Harvey, and R. G. Crystal. Threshold of biologic responses of the small airway epithelium to low levels of tobacco smoke. *American Journal of Respiratory and Critical Care Medicine*, 182(12):1524–1532, 2010.

[369] A. Tam, S. Wadsworth, D. Dorscheid, S.-F. P. Man, and D. D. Sin. Estradiol increases mucus synthesis in bronchial epithelial cells. *PLoS ONE*, 9(6):e100633, 2014.

[370] W. C. Tan, J. Bourbeau, S. D. Aaron, J. C. Hogg, F. Maltais, P. Hernandez, D. D. Marciniuk, K. R. Chapman, T. To, J. M. FitzGerald, B. L. Walker, J. Road, L. Zheng, G. Zhou, T. Yau, A. Benedetti, D. O'Donnell, and D. D. Sin. The effects of marijuana smoking on lung function in older people. *European Respiratory Journal*, 54(6):1900826, 2019.

[371] W. C. Tan, C. Lo, A. Jong, L. Xing, M. J. FitzGerald, W. M. Vollmer, S. A. Buist, and D. D. S. and. Marijuana and chronic obstructive lung disease: a population-based study. *Canadian Medical Association Journal*, 180(8):814–820, 2009.

[372] D. Tashkin. Cannabis smoking and the lung. In *Handbook of Cannabis and Related Pathologies*, pages 494–504. Elsevier, 2017.

[373] D. P. Tashkin. Effects of marijuana smoking on the lung. *Annals of the American Thoracic Society*, 10(3):239–247, 2013.

[374] D. P. Tashkin, A. H. Coulson, V. A. Clark, M. Simmons, L. B. Bourque, S. Duann, G. H. Spivey, and H. Gong. Respiratory symptoms and lung function in habitual heavy smokers of marijuana alone, smokers of marijuana and tobacco, smokers of tobacco alone, and nonsmokers. *American Review of Respiratory Disease*, 135(1):209–216, 1987.

[375] D. P. Tashkin and M. D. Roth. Pulmonary effects of inhaled cannabis smoke. *The American Journal of Drug and Alcohol Abuse*, 45(6):596–609, 2019. PMID: 31298945.

[376] D. P. Tashkin, B. J. Shapiro, and I. M. Frank. Acute pulmonary physiologic effects of smoked marijuana and oral δ9-tetrahydrocannabinol in healthy young men. *New England Journal of Medicine*, 289(7):336–341, 1973.

[377] D. P. Tashkin, B. J. Shapiro, and I. M. Frank. Acute effects of smoked marijuana and oral δ9-tetrahydrocannabinol on specific airway conductance in asthmatic subjects. *American Review of Respiratory Disease*, 109(4):420–428, 1974.

[378] D. P. Tashkin, B. J. Shapiro, Y. E. Lee, and C. E. Harper. Effects of smoked marijuana in experimentally induced asthma. *American Review of Respiratory Disease*, 112(3):377–386, 1975.

[379] D. P. Tashkin, B. J. Shapiro, Y. E. Lee, and C. E. Harper. Subacute effects of heavy marihuana smoking on pulmonary function in healthy men. *New England Journal of Medicine*, 294(3):125–129, 1976.

[380] S. A. Terlouw, R. Masereeuw, P. H. H. V. D. Broek, S. Notenboom, and F. G. M. Russel. Role of multidrug resistance protein 2 (MRP2) in glutathione-bimane efflux from caco-2 and rat renal proximal tubule cells. *British Journal of Pharmacology*, 134(5):931–938, 2001.

[381] M. J. Thun, B. D. Carter, D. Feskanich, N. D. Freedman, R. Prentice, A. D. Lopez, P. Hartge, and S. M. Gapstur. 50-year trends in smoking-related mortality in the united states. *New England Journal of Medicine*, 368(4):351–364, 2013.

[382] A. E. Tilley, B.-G. Harvey, A. Heguy, N. R. Hackett, R. Wang, T. P. O'Connor, and R. G. Crystal. Down-regulation of the notch pathway in human airway epithelium in association with smoking and chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 179(6):457–466, 2009.

[383] A. E. Tilley, T. P. O'Connor, N. R. Hackett, Y. Strulovici-Barel, J. Salit, N. Amoroso, X. K. Zhou, T. Raman, L. Omberg, A. Clark, J. Mezey, and R. G. Crystal. Biologic phenotyping of the human small airway epithelial response to cigarette smoking. *PLoS ONE*, 6(7):e22798, 2011.

[384] S. Trouillet-Assant, S. Viel, A. Gaymard, S. Pons, J.-C. Richard, M. Perret, M. Villard, K. Brengel-Pesce, B. Lina, M. Mezidi, L. Bitker, A. Belot, W. Mouton, G. Oriol, C. Compagnon, L. Generenaz, V. Cheynet, F. Ader, A. Becker, N. Benech, P. Chauvelot, C. Chidiac, A. Conrad, T. Ferry, P. Miailhes, T. Perpoint, M. Perry, C. Pouderoux, S. Roux, C. Triffault-Fillit, F. Valour, Y. Hodane, L. Chauvelot, P. Chabert, J. Provoost, G. David, L. Folliet, P. Lecam, G. Billaud, M. Bouscambert, V. Escuret, E. Frobert, A. Bal, G. Destras, L. Josset, F. Morfin, C. Munier, M. Valette, F. Venet, L. Garnier, R. Pescarmona, C. Lombard, and T. Walzer. Type i IFN immunoprofiling in COVID-19 patients. *Journal of Allergy and Clinical Immunology*, 146(1):206–208.e2, 2020.

[385] C.-T. K. Tseng, J. Tseng, L. Perrone, M. Worthy, V. Popov, and C. J. Peters. Apical entry and release of severe acute respiratory syndrome-associated coronavirus in polarized calu-3 lung epithelial cells. *Journal of Virology*, 79(15):9470–9479, 2005.

[386] C. Turcotte, M.-R. Blanchet, M. Laviolette, and N. Flamand. Impact of cannabis, cannabinoids, and endocannabinoids in the lungs. *Frontiers in Pharmacology*, 7, 2016.

[387] M. L. Turetz, T. P. O'Connor, A. E. Tilley, Y. Strulovici-Barel, J. Salit, D. Dang, M. Teater, J. Mezey, A. G. Clark, and R. G. Crystal. Trachea epithelium as a "canary" for cigarette smoking-induced biologic phenotype of the small airway epithelium. *Clinical and Translational Science*, 2(4):260–272, 2009.

[388] J. K. Tyzack, X. Wang, G. J. Belsham, and C. G. Proud. ABC50 interacts with eukaryotic initiation factor 2 and associates with the ribosome in an ATP-dependent manner. *Journal of Biological Chemistry*, 275(44):34131–34139, 2000.

[389] M. Uhlen, A. Bandrowski, S. Carr, A. Edwards, J. Ellenberg, E. Lundberg, D. L. Rimm, H. Rodriguez, T. Hiltke, M. Snyder, and T. Yamamoto. A proposal for validation of antibodies. *Nature Methods*, 13(10):823–827, 2016.

[390] L. Unterholzner, S. E. Keating, M. Baran, K. A. Horan, S. B. Jensen, S. Sharma, C. M. Sirois, T. Jin, E. Latz, T. S. Xiao, K. A. Fitzgerald, S. R. Paludan, and A. G. Bowie. IFI16 is an innate immune sensor for intracellular DNA. *Nature Immunology*, 11(11):997–1004, 2010.

[391] L. Vachon, M. X. FitzGerald, N. H. Solliday, I. A. Gould, and E. A. Gaensler. Single-dose effect of marihuana smoke. *New England Journal of Medicine*, 288(19):985–989, 1973.

[392] J. van Delft, S. Gaj, M. Lienhard, M. W. Albrecht, A. Kirpiy, K. Brauers, S. Claessen, D. Lizarraga, H. Lehrach, R. Herwig, and J. Kleinjans. RNA-seq provides new insights in the transcriptome responses induced by the carcinogen benzo[a]pyrene. *Toxicological Sciences*, 130(2):427–439, aug 2012.

[393] M. van der Deen, E. G. de Vries, W. Timens, R. J. Scheper, H. Timmer-Bosscha, and D. S. Postma. ATP-binding cassette (ABC) transporters in normal and pathological lung. *Respiratory Research*, 6(1), 2005.

[394] N. Van Doremalen, T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J. L. Harcourt, N. J. Thornburg, S. I. Gerber, et al. Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. *New England journal of medicine*, 382(16):1564–1567, 2020.

[395] V. Vasiliou, K. Vasiliou, and D. W. Nebert. Human ATP-binding cassette (ABC) transporter family. *Human Genomics*, 3(3):281, 2008.

[396] G. Veit, R. G. Avramescu, A. N. Chiang, S. A. Houck, Z. Cai, K. W. Peters, J. S. Hong, H. B. Pollard, W. B. Guggino, W. E. Balch, W. R. Skach, G. R. Cutting, R. A. Frizzell, D. N. Sheppard, D. M. Cyr, E. J. Sorscher, J. L. Brodsky, and G. L. Lukacs. From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. *Molecular Biology of the Cell*, 27(3):424–433, 2016.

[397] P. D. Vermeer, J. McHugh, T. Rokhlina, D. W. Vermeer, J. Zabner, and M. J. Welsh. Vaccinia virus entry, exit, and interaction with differentiated human airway epithelia. *Journal of Virology*, 81(18):9891–9899, 2007.

[398] J. Vestbo, S. S. Hurd, A. G. Agustí, P. W. Jones, C. Vogelmeier, A. Anzueto, P. J. Barnes, L. M. Fabbri, F. J. Martinez, M. Nishimura, R. A. Stockley, D. D. Sin, and R. Rodriguez-Roisin. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 187(4):347–365, 2013.

[399] C. Viniol and C. F. Vogelmeier. Exacerbations of copd. *European Respiratory Review*, 27(147), 2018.

[400] C. Vogel and E. M. Marcotte. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, 13(4):227–232, 2012.

[401] A. H. Wagener, C. Y. Yick, P. Brinkman, M. P. van der Schee, N. Fens, and P. J. Sterk. Toward composite molecular signatures in the phenotyping of asthma. *Annals of the American Thoracic Society*, 10(Supplement):S197–S205, 2013.

[402] M. S. Walters, B. P. De, J. Salit, L. J. Buro-Auriemma, T. Wilson, A. M. Rogalski, L. Lief, N. R. Hackett, M. R. Staudt, A. E. Tilley, B.-G. Harvey, R. J. Kaner, J. G. Mezey, B. Ashbridge, M. A. S. Moore, and R. G. Crystal. Smoking accelerates aging of the small airway epithelium. *Respiratory Research*, 15(1), 2014.

[403] M. S. Walters, J. Salit, J. H. Ju, M. R. Staudt, R. J. Kaner, A. M. Rogalski, T. B. Sodeinde, R. Rahim, Y. Strulovici-Barel, J. G. Mezey, A. M. Almulla, H. Sattar, M. Mahmoud, and R. G. Crystal. Waterpipe smoking induces epigenetic changes in the small airway epithelium. *PLOS ONE*, 12(3):e0171112, 2017.

[404] B. Wang, A. M. Mezlini, F. Demir, M. Fiume, Z. Tu, M. Brudno, B. Haibe-Kains, and A. Goldenberg. Similarity network fusion for aggregating data types on a genomic scale. *Nature methods*, 11(3):333–337, 2014.

[405] G. Wang, R. Wang, B. Ferris, J. Salit, Y. Strulovici-Barel, N. R. Hackett, and R. G. Crystal. Smoking-mediated up-regulation of GAD67 expression in the human airway epithelium. *Respiratory Research*, 11(1), 2010.

[406] G. Wang, Z. Xu, R. Wang, M. Al-Hijji, J. Salit, Y. Strulovici-Barel, A. E. Tilley, J. G. Mezey, and R. G. Crystal. Genes associated with MUC5ac expression in small airway epithelium of human smokers and non-smokers. *BMC Medical Genomics*, 5(1), 2012.

[407] K. Wang, W. Chen, Y.-S. Zhou, J.-Q. Lian, Z. Zhang, P. Du, L. Gong, Y. Zhang, H.-Y. Cui, J.-J. Geng, B. Wang, X.-X. Sun, C.-F. Wang, X. Yang, P. Lin, Y.-Q. Deng, D. Wei, X.-M. Yang, Y.-M. Zhu, K. Zhang, Z.-H. Zheng, J.-L. Miao, T. Guo, Y. Shi, J. Zhang, L. Fu, Q.-Y. Wang, H. Bian, P. Zhu, and Z.-N. Chen. *SARS-CoV-2 invades host cells via a novel route: CD147-spike protein*. Cold Spring Harbor Laboratory, 2020.

[408] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA*, 2020.

[409] Z. Wang, M. Gerstein, and M. Snyder. RNA-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63, jan 2009.

[410] K. Watanabe, S. Yamaori, T. Funahashi, T. Kimura, and I. Yamamoto. Cytochrome p450 enzymes involved in the metabolism of tetrahydrocannabinols and cannabinol by human hepatic microsomes. *Life Sciences*, 80(15):1415–1419, 2007.

[411] H. wei Jiang, H. nan Zhang, Q. feng Meng, J. Xie, Y. Li, H. Chen, Y. xiao Zheng, X. ning Wang, H. Qi, J. Zhang, P.-H. Wang, Z.-G. Han, and S. ce Tao. SARS-CoV-2 orf9b suppresses type i interferon responses by targeting TOM70. *Cellular & Molecular Immunology*, 17(9):998–1000, 2020.

[412] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, 2009.

[413] S. M. Wilcox, H. Arora, L. Munro, J. Xin, F. Fenninger, L. A. Johnson, C. G. Pfeifer, K. B. Choi, J. Hou, P. A. Hoodless, and W. A. Jefferies. The role of the innate immune response regulatory gene ABCF1 in mammalian embryogenesis and development. *PLOS ONE*, 12(5):e0175918, 2017.

[414] A. J. Wilk, A. Rustagi, N. Q. Zhao, J. Roque, G. J. Martínez-Colón, J. L. McKechnie, G. T. Ivison, T. Ranganath, R. Vergara, T. Hollis, L. J. Simpson, P. Grant, A. Subramanian, A. J. Rogers, and C. A. Blish. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nature Medicine*, 26(7):1070–1076, 2020.

[415] H. Wirtz and M. Schmidt. Acute influence of cigarette smoke on secretion of pulmonary surfactant in rat alveolar type II cells in culture. *European Respiratory Journal*, 9(1):24–32, 1996.

[416] P. G. Woodruff, H. A. Boushey, G. M. Dolganov, C. S. Barker, Y. H. Yang, S. Donnelly, A. Ellwanger, S. S. Sidhu, T. P. Dao-Pick, C. Pantoja, D. J. Erle, K. R. Yamamoto, and J. V. Fahy. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proceedings of the National Academy of Sciences*, 104(40):15858–15863, 2007.

[417] P. G. Woodruff, B. Modrek, D. F. Choy, G. Jia, A. R. Abbas, A. Ellwanger, J. R. Arron, L. L. Koth, and J. V. Fahy. T-helper type 2–driven inflammation defines major subphenotypes of asthma. *American journal of respiratory and critical care medicine*, 180(5):388–395, 2009.

[418] P. G. Woodruff, B. Modrek, D. F. Choy, G. Jia, A. R. Abbas, A. Ellwanger, J. R. Arron, L. L. Koth, and J. V. Fahy. T-helper type 2–driven inflammation defines major subphenotypes of asthma. *American Journal of Respiratory and Critical Care Medicine*, 180(5):388–395, 2009.

[419] D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, and J. S. McLellan. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483):1260–1263, 2020.

[420] T.-C. WU, D. P. Tashkin, B. Djahed, and J. E. Rose. Pulmonary hazards of smoking marijuana as compared with tobacco. *New England Journal of Medicine*, 318(6):347–351, 1988.

[421] R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, and C. Wendtner. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 581(7809):465–469, 2020.

[422] H. Xia, Z. Cao, X. Xie, X. Zhang, J. Y.-C. Chen, H. Wang, V. D. Menachery, R. Rajsbaum, and P.-Y. Shi. Evasion of type i interferon by SARS-CoV-2. *Cell Reports*, 33(1):108234, 2020.

[423] Y. Xing, J. Chen, J. Tu, B. Zhang, X. Chen, H. Shi, S. C. Baker, L. Feng, and Z. Chen. The papain-like protease of porcine epidemic diarrhea virus negatively regulates type i interferon pathway by acting as a viral deubiquitinase. *Journal of General Virology*, 94(7):1554–1567, 2013.

[424] K. Xu, H. Cai, Y. Shen, Q. Ni, Y. Chen, S. Hu, J. Li, H. Wang, L. Yu, H. Huang, Y. Qiu, G. Wei, Q. Fang, J. Zhou, J. Sheng, T. Liang, and L. Li. Translation: Management of coronavirus disease 2019 (COVID-19): Experience in zhejiang province, china. *Infectious Microbes and Diseases*, 2(2):55–63, 2020.

[425] S. Yamakawa and K. Hayashida. Advances in surgical applications of growth factors for wound healing. *Burns & Trauma*, 7, 2019.

[426] N. Yan, P. Cherepanov, J. E. Daigle, A. Engelman, and J. Lieberman. The SET complex acts as a barrier to autointegration of HIV-1. *PLoS Pathogens*, 5(3):e1000327, 2009.

[427] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, 367(6485):1444–1448, 2020.

[428] J. Yang, W.-L. Zuo, T. Fukui, I. Chao, K. Gomi, B. Lee, M. R. Staudt, R. J. Kaner, Y. Strulovici-Barel, J. Salit, R. G. Crystal, and R. Shaykhiev. Smoking-dependent distal-to-proximal repatterning of the adult human small airway epithelium. *American Journal of Respiratory and Critical Care Medicine*, 196(3):340–352, 2017.

[429] Y. Yang, L. Zhang, H. Geng, Y. Deng, B. Huang, Y. Guo, Z. Zhao, and W. Tan. The structural and accessory proteins m, ORF 4a, ORF 4b, and ORF 5 of middle east respiratory syndrome coronavirus (MERS-CoV) are potent interferon antagonists. *Protein & Cell*, 4(12):951–961, 2013.

[430] X. Yin, L. Riva, Y. Pu, L. Martin-Sancho, J. Kanamune, Y. Yamamoto, K. Sakai, S. Gotoh, L. Miorin, P. D. De Jesus, et al. Mda5 governs the innate immune response to sars-cov-2 in lung epithelial cells. *Cell reports*, 34(2):108628, 2021.

[431] G. T. Yocum, J. Chen, C. H. Choi, E. A. Townsend, Y. Zhang, D. Xu, X. W. Fu, M. J. Sanderson, and C. W. Emala. Role of transient receptor potential vanilloid 1 in the modulation of airway smooth muscle tone and calcium handling. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 312(6):L812–L821, 2017.

[432] T. Zaman, T. Moua, E. Vittinghoff, J. H. Ryu, H. R. Collard, and J. S. Lee. Differences in clinical characteristics and outcomes between men and women with idiopathic pulmonary fibrosis: a multicenter retrospective cohort study. *Chest*, 158(1):245–251, 2020.

[433] H. Zeng, G. Liu, P. A. Rea, and G. D. Kruh. Transport of amphipathic anions by human multidrug resistance protein 3. *Cancer Research*, 60(17):4779–4784, 2000.

[434] L. R. Zhang, H. Morgenstern, S. Greenland, S.-C. Chang, P. Lazarus, M. D. Teare, P. J. Woll, I. Orlow, B. Cox, Y. Brhane, G. Liu, and R. J. H. and. Cannabis smoking and lung cancer risk: Pooled analysis in the international lung cancer consortium. *International Journal of Cancer*, 136(4):894–903, 2014.

[435] Q. Zhang, P. Bastard, Z. Liu, J. Le Pen, M. Moncada-Velez, J. Chen, M. Ogishi, I. K. Sabli, S. Hodeib, C. Korol, et al. Inborn errors of type i ifn immunity in patients with life-threatening covid-19. *Science*, 370(6515), 2020.

[436] R. Zhang, T.-K. Kim, Z.-H. Qiao, J. Cai, W. M. Pierce, and Z.-H. Song. Biochemical and mass spectrometric characterization of the human CB2 cannabinoid receptor expressed in pichia pastoris—importance of correct processing of the n-terminus. *Protein Expression and Purification*, 55(2):225–235, 2007.

[437] Y. Zhang, R. Liu, M. Ni, P. Gill, and A. S. Lee. Cell surface relocalization of the endoplasmic reticulum chaperone and unfolded protein response regulator GRP78/BiP. *Journal of Biological Chemistry*, 285(20):15065–15075, 2010.

[438] S. Zhao, W.-P. Fung-Leung, A. Bittner, K. Ngo, and X. Liu. Comparison of rna-seq and microarray in transcriptome profiling of activated t cells. *PloS one*, 9(1):e78644, 2014.

[439] H. Zhou, A. Brekman, W.-L. Zuo, X. Ou, R. Shaykhiev, F. J. Agosto-Perez, R. Wang, M. S. Walters, J. Salit, Y. Strulovici-Barel, M. R. Staudt, R. J. Kaner, J. G. Mezey, R. G. Crystal, and G. Wang. POU2af1 functions in the human airway epithelium to regulate expression of host defense genes. *The Journal of Immunology*, 196(7):3159–3167, 2016.

[440] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature*, 579(7798):270–273, 2020.

[441] Z. Zhou, L. Ren, L. Zhang, J. Zhong, Y. Xiao, Z. Jia, L. Guo, J. Yang, C. Wang, S. Jiang, D. Yang, G. Zhang, H. Li, F. Chen, Y. Xu, M. Chen, Z. Gao, J. Yang, J. Dong, B. Liu, X. Zhang, W. Wang, K. He, Q. Jin, M. Li, and J. Wang. Heightened innate immune responses in the respiratory tract of COVID-19 patients. *Cell Host & Microbe*, 27(6):883–890.e2, 2020.

[442] J. Zhu, S. D. Message, P. Mallia, T. Kebadze, M. Contoli, C. K. Ward, E. S. Barnathan, M. A. Mascelli, O. M. Kon, A. Papi, et al. Bronchial mucosal ifn-α/β and pattern recognition receptor expression in patients with experimental rhinovirus-induced asthma exacerbations. *Journal of Allergy and Clinical Immunology*, 143(1):114–125, 2019.

[443] Z. Zhu, X. Lian, X. Su, W. Wu, G. A. Marraro, and Y. Zeng. From sars and mers to covid-19: a brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respiratory research*, 21(1):1–14, 2020.

[444] C. Ziegler, S. J. Allon, S. K. Nyquist, I. Mbano, V. N. Miao, Y. Cao, A. S. Yousif, J. Bals, B. M. Hauser, J. Feldman, C. Muus, M. H. W. II, S. Kazer, T. K. Hughes, B. Doran, G. J. Gatter, M. Vukovic, C. N. Tzouanas, F. Taliaferro, Z. Guo, J. P. Wang, D. F. Dwyer, K. M. Buchheit, J. Boyce, N. A. Barrett, T. M. Laidlaw, S. L. Carroll, L. Colonna, V. Tkachev, A. Yu, H. B. Zheng, H. P. Gideon, C. G. Winchell, P. L. Lin, B. Berger, A. Leslie, J. L. Flynn, S. M. Fortune, R. W. Finberg, L. Kean, M. Garber, A. Schmidt, D. Lingwood, A. K. Shalek, J. Ordovas-Montanes, and H. L. B. Network. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is enriched in specific cell subsets across tissues. *SSRN Electronic Journal*, 2020.

[445] L. Zou, F. Ruan, M. Huang, L. Liang, H. Huang, Z. Hong, J. Yu, M. Kang, Y. Song, J. Xia, Q. Guo, T. Song, J. He, H.-L. Yen, M. Peiris, and J. Wu. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *New England Journal of Medicine*, 382(12):1177–1179, 2020.

[446] P. M. Zygmunt, J. Petersson, D. A. Andersson, H. hu Chuang, M. Sørgård, V. D. Marzo, D. Julius, and E. D. Högestätt. Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide. *Nature*, 400(6743):452–457, 1999.

APPENDICES

Appendix A
Supplementary Material - Chapter 2

Supplementary Table 1

Table 1: GEO data set demographics

Comparison of Dataset Demographics									
GEO ID	GSE994	GSE4498	GSE11784	GSE11906	GSE37147	GSE4302	GSE63142	GSE67472	GSE76227
Airway Generation	Large airway (generation 2^{nd}-3^{rd})	Small airway (generation 10^{th}-12^{th})	Small airway (generation 10^{th}-12^{th})	Trachea, large (generation 2^{nd}-3^{rd}), & small airway (generation 10^{th}-12^{th})	Medium airway (generation 6^{th}-8^{th})	Medium airway (generation 3^{rd}-5^{th})	Medium airway (generation 3^{rd}-5^{th})	Medium airway (generation 3^{rd}-5^{th})	Medium airway (generation 3^{rd}-5^{th})
Sample Type	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings	Epithelial cells from brushings
Number and Classification of Patients	34 current smokers, 14 former smokers (average duration of smoking cessation was 10.49 years), and 23 never-smokers, with current and former smokers having >22 pack years smoking history with no reported COPD	10 individuals with >25 pack years smoking history with no reported COPD	72 individuals with >25 pack years smoking history with no reported COPD 36 individuals with >34 pack years smoking history with reported COPD, 72 individuals with >25 pack years smoking history with no reported COPD	17 trachea, 21 large airway, and 35 small airway from healthy individuals 54 individuals with >25 pack years smoking history with no reported COPD 20 individuals with >38 pack years smoking history with reported COPD, 54 individuals with >25 pack years smoking history with no reported COPD	69 current smokers and 82 former smokers (average duration of smoking cessation was 11.11 years) with >47 pack years smoking history with no reported COPD 87 individuals with >51 pack years smoking history with reported COPD, 151 individuals with >47 pack years smoking history with no reported COPD	28 healthy individuals and 42 asthmatics	26 healthy individuals, 59 mild asthmatics, 19 moderate asthmatics and 51 severe asthmatics	43 healthy individuals and 62 asthmatics	26 healthy individuals, 59 mild asthmatics, 19 moderate asthmatics and 51 severe asthmatics
Microarray Platform	Affymetrix Human Genome U133A	Affymetrix Human Genome U133 Plus 2	Affymetrix Human Genome U133 Plus 2	Affymetrix Human Genome U133 Plus 2	Affymetrix Human Gene 1 ST	Affymetrix Human Genome U133 Plus 2	Agilent 014850 Whole Human Genome Microarray 4x44K G4112F	Affymetrix Human Genome U133 Plus 2	Affymetrix HT HG U133 plus PM
Normalization Method	Mas5 normalization without log transformation	Mas5 normalization without log transformation	Mas5 normalization without log transformation	Mas5 normalization without log transformation	Robust Multi-array Average (RMA) with log transformation	Robust Multi-array Average (RMA) with log transformation	Cyclic LOESS	Robust Multi-array Average (RMA) with log transformation	Robust Multi-array Average (RMA) with log transformation

Supplementary Table 2

Table 2: ABC transporter expression in different airway generations

GSE11906 – Small vs Large Airway				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Small Airway
230913_at	ABCG1	0.180	0.180	-0.894
202804_at	ABCC1	0.180	0.180	-0.874
1557374_at	ABCC9	0.180	0.180	-0.851
217504_at	ABCA6	0.180	0.180	-0.776
213485_s_at	ABCC10	0.180	0.180	-0.769
243951_at	ABCB1	0.180	0.180	-0.693
203196_at	ABCC4	0.180	0.180	-0.681
1555323_at	ABCB9	0.180	0.180	-0.631
206155_at	ABCC2	0.180	0.180	-0.614
203192_at	ABCB6	0.180	0.180	-0.613

GSE11906 – Large Airway vs Trachea				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Large Airway
204719_at	ABCA8	0.053	0.079	0.963
1552590_a_at	ABCC12	0.053	0.079	0.854
220383_at	ABCG5	0.053	0.079	0.807
1570505_at	ABCB4	0.053	0.079	0.347
1554911_at	ABCC11	0.053	0.079	0.133
209735_at	ABCG2	0.074	0.110	-0.686
233371_at	ABCC13	0.074	0.110	0.423
240717_at	ABCB5	0.037	0.111	0.912
210245_at	ABCC8	0.101	0.152	0.389
231751_at	ABCG8	0.136	0.204	1.474

GSE11906 – Small Airway vs Trachea				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Small Airway
202804_at	ABCC1	0.007	0.022	-0.983
203196_at	ABCC4	0.007	0.022	-0.916
225973_at	ABCB3	0.007	0.022	-0.902
230913_at	ABCG1	0.007	0.022	-0.866
1557374_at	ABCC9	0.007	0.022	-0.849
213485_s_at	ABCC10	0.007	0.022	-0.752
200045_at	ABCF1	0.007	0.022	-0.743
219577_s_at	ABCA7	0.007	0.022	-0.678
203192_at	ABCB6	0.007	0.022	-0.627
1554878_at	ABCD3	0.007	0.022	-0.573

Supplementary Table 3

Table 3: Impact of smoking status on ABC transporter expression

GSE11906 - Non-smokers vs Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Smokers
203192_at	**ABCB6**	1.58E-09	1.79E-07	1.194
208161_s_at	**ABCC3**	7.65E-08	4.32E-06	1.045
1553295_at	ABCA13	1.58E-06	4.45E-05	-1.045
1553604_at	**ABCA13**	1.20E-06	4.45E-05	-0.731
209620_s_at	ABCB7	5.58E-06	1.26E-04	-0.455
1553605_a_at	ABCA13	8.82E-06	1.58E-04	-0.739
202804_at	**ABCC1**	9.76E-06	1.58E-04	0.562
209641_s_at	ABCC3	5.51E-05	7.79E-04	1.172
230682_x_at	ABCC3	7.25E-05	9.11E-04	0.636
223320_s_at	ABCB10	1.19E-04	1.22E-03	-0.219

GSE11784 - Non-smokers vs Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Smokers
203192_at	**ABCB6**	4.64E-10	5.24E-08	0.902
202804_at	**ABCC1**	2.75E-09	1.56E-07	0.624
208161_s_at	**ABCC3**	3.63E-06	1.37E-03	0.857
207622_s_at	ABCF2	1.83E-04	5.18E-03	-0.258
1553605_a_at	ABCA13	2.46E-04	5.55E-03	-0.785
230682_x_at	ABCC3	4.09E-04	7.71E-03	0.378
216066_at	ABCA1	2.84E-03	0.046	0.663
1553604_at	**ABCA13**	5.82E-03	0.072	-0.509
209620_s_at	ABCB7	5.34E-03	0.072	-0.212
211224_s_at	ABCB11	6.34E-03	0.072	0.390

GSE4498 - Non-smokers vs Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Smokers
203192_at	**ABCB6**	1.39E-04	0.016	0.928
208161_s_at	**ABCC3**	6.00E-04	0.034	1.006
1553604_at	**ABCA13**	5.64E-03	0.213	-1.124
1553605_a_at	ABCA13	0.014	0.311	-1.022
215559_at	ABCC6	0.021	0.388	-0.322
1558460_at	ABCC5	0.030	0.405	-0.293
209994_s_at	ABCB1	0.030	0.405	0.271
1569072_s_at	ABCB5	0.036	0.405	0.204
202804_at	**ABCC1**	0.036	0.405	0.296
209641_s_at	ABCC3	0.043	0.437	0.377

Supplementary Table 4

Table 4: Impact of smoking cessation on ABC transporter expression

GSE37147 - Current vs Former Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Smokers
85320_at	ABCC11	3.89E-04	0.019	0.002
24_at	ABCC4	1.06E-03	0.026	-0.007
8714_at	**ABCC3**	2.79E-03	0.041	0.137
10058_at	**ABCB6**	3.31E-03	0.041	0.139
22_at	ABCB7	5.22E-03	0.045	-0.012
5244_at	ABCB4	5.47E-03	0.045	0.042
4363_at	**ABCC1**	6.89E-03	0.046	0.069
5825_at	ABCD3	7.47E-03	0.046	-0.011
154664_at	**ABCA13**	0.013	0.067	-0.058
64137_at	ABCG4	0.014	0.067	-0.010

GSE994 - Current vs Former Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Current Smokers
202804_at	**ABCC1**	1.52E-04	2.29E-04	0.430
202805_s_at	ABCC1	5.45E-04	1.63E-03	0.564
203192_at	**ABCB6**	1.92E-03	5.75E-03	0.401
215703_at	CFTR	1.93E-03	5.80E-03	0.343
215873_x_at	ABCC10	1.96E-03	5.88E-03	-0.042
219577_s_at	ABCA7	0.014	0.021	-0.529
209641_s_at	**ABCC3**	9.77E-03	0.029	0.160
209993_at	ABCB1	0.042	0.064	-0.388
205043_at	CFTR	0.024	0.073	0.156
201873_s_at	ABCE1	0.026	0.078	0.119

GSE994 - Current vs Never Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Current Smokers
202804_at	**ABCC1**	9.68E-05	2.90E-04	0.603
219577_s_at	ABCA7	8.10E-04	2.43E-03	-0.490
203981_s_at	ABCD4	7.02E-03	0.021	-0.317
203192_at	**ABCB6**	0.016	0.024	0.604
202394_s_at	ABCF3	0.020	0.061	-0.172
201873_s_at	ABCE1	0.053	0.080	0.268
203982_s_at	ABCD4	0.027	0.081	-0.510
210245_at	ABCC8	0.029	0.088	-0.193
215559_at	ABCC6	0.031	0.093	-0.503
202805_s_at	ABCC1	0.063	0.094	0.633

GSE994 - Former vs Never Smokers				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Former Smokers
215703_at	CFTR	9.08E-03	0.014	0.899
205043_at	CFTR	0.034	0.051	-0.792
203196_at	ABCC4	0.023	0.068	-0.301
215465_at	ABCA12	0.055	0.082	0.475
202850_at	ABCD3	0.028	0.084	-0.128
214033_at	ABCC6	0.063	0.095	-0.294
202805_s_at	ABCC1	0.109	0.109	0.069
206155_at	ABCC2	0.041	0.122	-0.394
215873_x_at	ABCC10	0.158	0.158	-0.089
208561_at	ABCC9	0.060	0.180	-0.778

Supplementary Table 5

Table 5: Association between COPD status and ABC transporter expression

GSE11906 - Smoker without COPD vs Smoker with COPD				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in COPD
203505_at	ABCA1	1.91E-04	0.011	0.595
1555039_a_at	ABCC4	1.43E-04	0.011	-0.780
207583_at	ABCD2	2.59E-03	0.057	-0.493
1554878_a_at	ABCD3	2.39E-03	0.057	-0.551
204567_s_at	ABCG1	2.12E-03	0.057	0.385
232081_at	ABCG1	3.04E-03	0.057	0.457
202804_at	**ABCC1**	7.28E-03	0.091	0.262
1554918_a_at	ABCC4	7.28E-03	0.091	-0.494
1558460_at	ABCC5	6.52E-03	0.091	-0.283
204343_at	ABCA3	0.018	0.154	-0.485

GSE11784 - Smoker without COPD vs Smoker with COPD				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in COPD
213485_s_at	ABCC10	1.09E-03	0.123	0.338
226363_at	ABCC5	4.85E-03	0.274	-0.605
209994_s_at	ABCB1	0.028	0.344	-2.124
203191_at	ABCB6	0.015	0.344	0.745
223320_s_at	ABCB10	0.013	0.344	-0.440
202804_at	**ABCC1**	0.030	0.344	0.208
209380_s_at	ABCC5	0.018	0.344	-0.265
208480_s_at	ABCC6	0.030	0.344	-1.113
202850_at	ABCD3	0.028	0.344	-0.267
232081_at	ABCG1	0.030	0.344	0.429

GSE37147 - Smoker without COPD vs Smoker with COPD				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in COPD
225_at	ABCD2	8.37E-07	2.87E-05	-0.0716
10347_at	ABCA7	1.17E-06	2.87E-05	0.0321
26154_at	ABCA12	2.10E-05	2.62E-04	0.132
6891_at	TAP2	2.14E-05	2.62E-04	0.0387
6890_at	TAP1	5.72E-04	5.60E-03	0.0406
4363_at	**ABCC1**	8.29E-04	6.77E-03	0.0323
19_at	ABCA1	1.65E-03	0.012	0.0549
23456_at	ABCB10	3.49E-03	0.020	-0.00849
89845_at	ABCC10	3.60E-03	0.020	0.0128
9429_at	ABCG2	0.029	0.141	-0.0314

Supplementary Table 6

Table 6: Association between asthma status and ABC transporter expression

GSE67472 - No Asthma vs Asthma				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Asthmatics
154664_at	**ABCA13**	3.19E-07	1.50E-05	-0.438
1244_at	**ABCC2**	8.18E-06	1.92E-04	0.258
10057_at	ABCC5	1.81E-04	2.84E-03	-0.219
10060_at	**ABCC9**	4.39E-04	5.16E-03	-0.196
64240_at	ABCG5	9.43E-04	8.86E-03	-0.237
225_at	ABCD2	3.00E-03	0.023	-0.223
10351_at	ABCA8	9.59E-03	0.064	-0.161
10257_at	**ABCC4**	0.012	0.071	-0.146
8714_at	ABCC3	0.020	0.105	-0.099
10350_at	ABCA9	0.042	0.196	-0.118

GSE4302 - No Asthma vs Asthma				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	Log2 FC in Asthmatics
206155_at	**ABCC2**	7.07E-05	7.99E-03	0.167
202805_s_at	**ABCC1**	6.88E-04	0.039	0.152
1553604_at	**ABCA13**	2.62E-03	0.079	-0.380
1553605_a_at	ABCA13	3.49E-03	0.079	-0.325
207623_at	ABCF2	3.22E-03	0.079	0.074
208462_s_at	**ABCC9**	5.00E-03	0.089	-0.167
210099_at	ABCA2	5.51E-03	0.089	0.129
1554918_a_at	**ABCC4**	0.010	0.128	-0.111
207583_at	ABCD2	0.010	0.128	-0.119
1557374_at	ABCC9	0.017	0.188	-0.506

Supplementary Table 7

Table 7: Association between asthma severity and ABC transporter expression

GSE63142 - Mild/Moderate Asthma vs Severe Asthma				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	FC in Less Severe vs More Severe
A_23_P160940	ABCA4	4.75E-04	0.012	-0.143
A_23_P25559	**ABCC4**	8.43E-04	0.012	-0.118
A_24_P16913	ABCC4	5.43E-04	0.012	-0.081
A_24_P222291	ABCC4	7.27E-04	0.012	-0.062
A_23_P41380	ABCE1	3.65E-04	0.012	-0.162
A_23_P18713	ABCG2	1.01E-03	0.012	-0.197
A_23_P201918	ABCB10	1.63E-03	0.017	-0.097
A_23_P39481	ABCA7	5.30E-03	0.048	0.206
A_23_P219013	ABCC10	5.83E-03	0.048	0.081
A_24_P197196	ABCB6	8.98E-03	0.067	0.100

GSE76227 - Moderate Asthma vs Severe Asthma				
Probeset ID	Probeset Name	Raw p values	Benjamini Hochberg Adjustment	FC in Less Severe vs More Severe
1553604_PM_at	**ABCA13**	6.04E-04	0.065	-0.304
1553605_PM_a_at	ABCA13	2.12E-03	0.114	-0.251
1553295_PM_at	ABCA13	4.02E-03	0.145	-0.211
209246_PM_at	ABCF2	8.92E-03	0.241	0.183
214209_PM_s_at	ABCB9	0.019	0.414	0.182
241705_PM_at	ABCA5	0.041	0.427	-0.226
240717_PM_at	ABCB5	0.044	0.427	0.091
243167_PM_at	ABCB5	0.043	0.427	-0.146
1554918_PM_a_at	**ABCC4**	0.035	0.427	-0.077
205142_PM_x_at	ABCD1	0.027	0.427	0.108

Appendix B
Supplementary Material - Chapter 3

Supplementary Figure 1

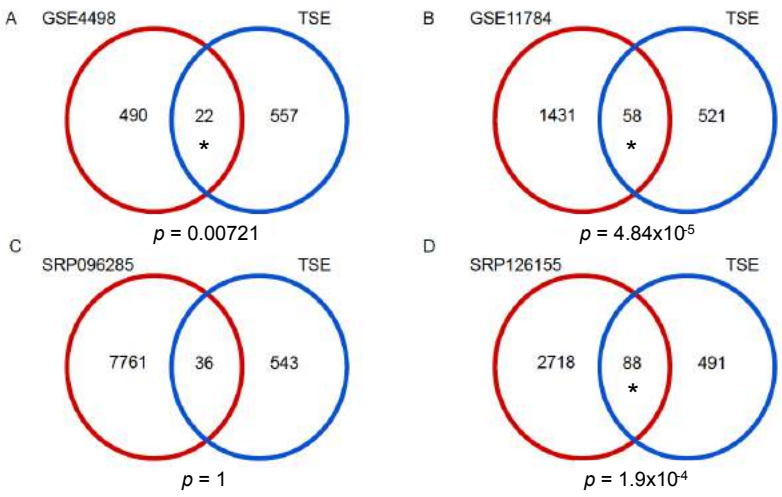


Figure 1: Union of differentially expressed genes between the tobacco smoke exposure experiment in Calu-3 cells from this study and the genes differentially expressed in (A) GSE4498, (B) GSE11784, (C) SRP096285, and (D) SRP126155 calculated with a hypergeometric test in R.

Supplementary Figure 2

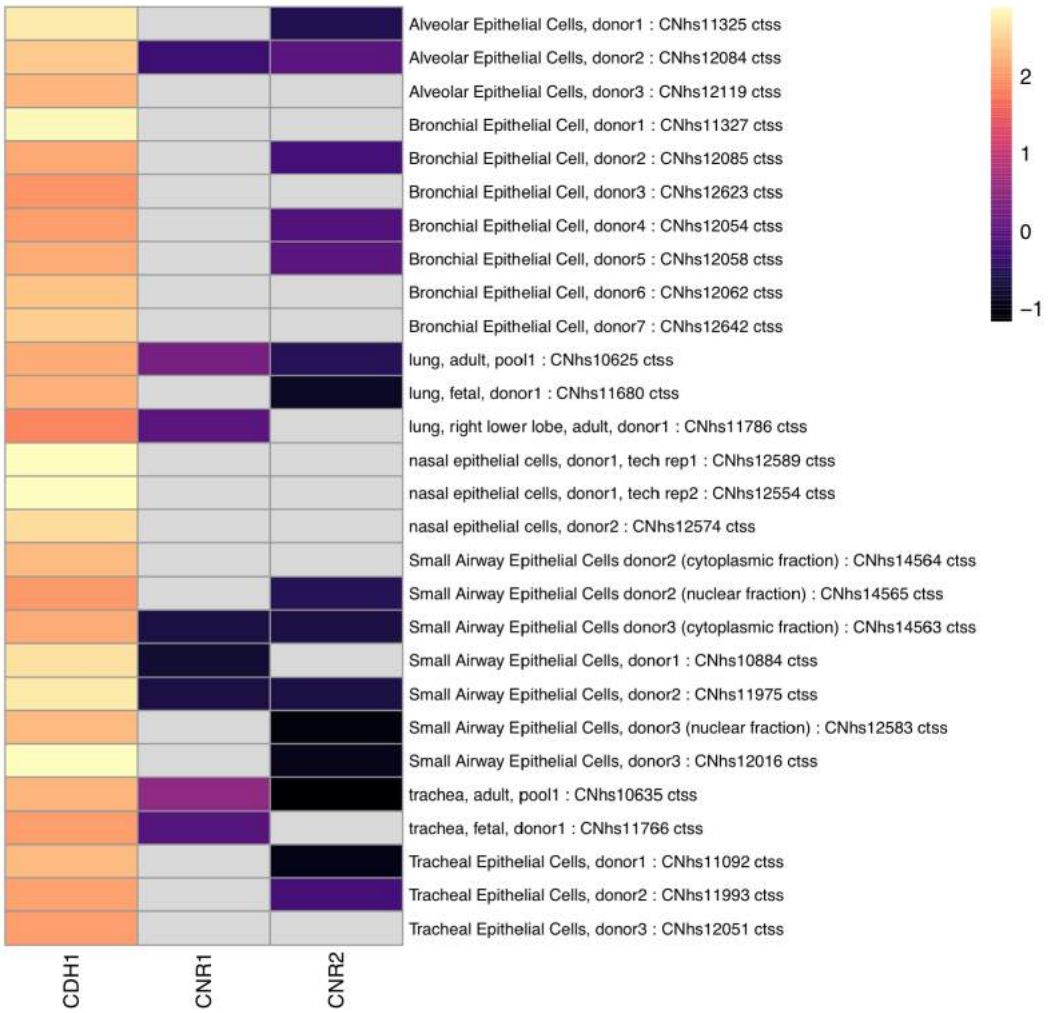


Figure 2: Heat map of FANTOM5 CAGE promoter activity data for CNR1, CNR2, and CDH1 (positive control) are shown for samples related to lung tissues ($n = 28$). Heat map colour is proportional to promoter activity, depicted as $\log_{10}$-transformed normalized transcripts per million (TPM). Grey represents no detection.

Supplementary Table 8

Table 8: Relevant endocannabinoid signalling pathway candidates

Molecule	Gene	Associated protein	Function
Receptor	*CNR1*	Cannabinoid receptor 1 (CB1)	Primary receptor involved in endocannabinoid signalling
	CNR2	Cannabinoid receptor 2 (CB2)	Primary receptor involved in endocannabinoid signalling
	GABRA2	γ-aminobutyric acid receptor subunit α-2 (GABRA2)	Implicated in cannabis dependency
	GPR55	G protein-coupled receptor 55 (GPR55)	Novel cannabinoid receptor
	OPRM1	Opioid receptor μ1 (MOR1)	Implicated in cannabis dependency
	TRPV1	Transient receptor potential vanilloid 1 (TRPV1)	Novel cannabinoid receptor
Enzyme	*ABHD12*	2-AG hydrolase ABHD12 (ABHD12)	Degradation of 2-AG
	ABHD6	2-AG hydrolase ABHD6 (ABHD6)	Degradation of 2-AG
	ADCY3	Adenylyl cyclase 3 (AC)	Catalyses formation of cAMP
	AKT1	AKT serine/threonine kinase 1 (AKT)	Regulates cell survival
	COMT	Catechol-*O*-methyltransferase (COMT)	Degradation of dopamine
	CYP2C9	Cytochrome P450 2C9 (CYP2C9)	Metabolism of THC
	CYP3A4	Cytochrome P450 3A4 (CYP3A4)	Metabolism of THC
	DAGLA	Diacylglycerol lipase α (DAGLA)	Biosynthesis of 2-AG
	DAGLB	Diacylglycerol lipase β (DAGLB)	Biosynthesis of 2-AG
	DUSP6	Dual-specificity phosphatase 6 (MKP3)	Regulates MAPK signalling
	FAAH	Fatty acid amide hydrolase (FAAH)	Degradation of AEA
	FAAH2	Fatty acid amide hydrolase 2 (FAAH2)	Degradation of AEA
	MAPK14	Mitogen-activated protein kinase 14 (MAPK14)	Regulates cell survival
	MAP2K2	Mitogen-activated protein kinase kinase 2 (MAP2K2)	Regulates cell survival
	MAPK3	Extracellular signal-regulated kinase (MAKP3)	Regulates cell survival
	MGLL	Monoglyceride lipase (MAGL)	Degradation of 2-AG
	NAAA	*N*-acylethanolamine acid amidase (NAAA)	Degradation of AEA
	NAPEPLD	*N*-acylphosphatidylethanolamine phospholipase D (NAPEPLD)	Biosynthesis of AEA
	NOS2	Inducible nitric oxide synthase (iNOS)	Inflammatory mediator
	PIK3CA	Phosphatidylinositol-3-kinase (PI3 K)	Regulates cell survival
	PRKACA	Protein kinase-A (PKA)	Regulates cell survival
	PTGS2	Cyclooxygenase-2 (COX2)	Inflammatory mediator
Other proteins	*ABCB1*	P-glycoprotein 1 (p-GP)	Cannabinoid transportation
	GNAI1	$G_{i/o}$ α subunit ($G_{i/o}$)	Coupled to cannabinoid receptors
	NRG1	Neuregulin 1 (NRG1)	Mediates cell–cell signalling
	TP53	Tumour protein p53 (p53)	Regulates cell survival

2-AG: 2-arachidonoylglycerol; THC: tetrahydrocannabinol; AEA: anandamide.

Supplementary Table 9

Table 9: GEO data sets analyzed

GSE accession	Affymetrix chip	Asthma	COPD	Healthy
GSE4302	HG-U133 Plus 2	74	0	13
GSE4498	HG-U133 Plus 2	0	0	12 (2F/10M)
GSE5058	HG-U133 Plus 2	0	14 (4F/10M)	0
GSE7832	HG-U133 Plus 2	0	0	8 (2F/6M)
GSE8545	HG-U133 Plus 2	0	15 (3F/12M)	5 (1F/4M)
GSE10006	HG-U133 Plus 2	0	20 (4F/16M)	21 (2F/19M)
GSE11784	HG-U133 Plus 2	0	17 (1F/3M)	40 (18F/22M)
GSE11906	HG-U133 Plus 2	0	0	30 (10F/20M)
GSE13931	HG-U133 Plus 2	0	0	19 (4F/15M)
GSE13933	HG-U133 Plus 2	0	0	11 (6F/5M)
GSE14224	HuEx-1.0-st-v2	0	0	11 (7F/4M)
GSE17905	HG-U133 Plus 2	0	0	1 (1M)
GSE19667	HG-U133 Plus 2	0	0	3 (3F)
GSE20257	HG-U133 Plus 2	0	1 (1F)	0
GSE22047	HG-U133 Plus 2	0	23	81
GSE34450	HG-U133 Plus 2	0	0	11
GSE37147	HuGene-1.0-st-v1	0	110 (35F/52M)	8
GSE40364	HG-U133 Plus 2	0	7	0
GSE43079	HG-U133 Plus 2	0	0	16
GSE43939	HG-U133 Plus 2	0	0	13
GSE52237	HG-U133 Plus 2	0	0	2
GSE64614	HG-U133 Plus 2	0	0	31
GSE67472	HG-U133 Plus 2	62 (34F/28M)	0	43 (20F/23M)
GSE77658	HG-U133 Plus 2	0	0	6
GSE84101	HG-U133 Plus 2	0	0	7
GSE97010	HuGene-1.0-st-v1	0	0	126 (28F/98M)
GSE108134	HG-U133 Plus 2	0	131	98

F: female; M: male

Appendix C
Supplementary Material - Chapter 4

Supplementary Figure 3

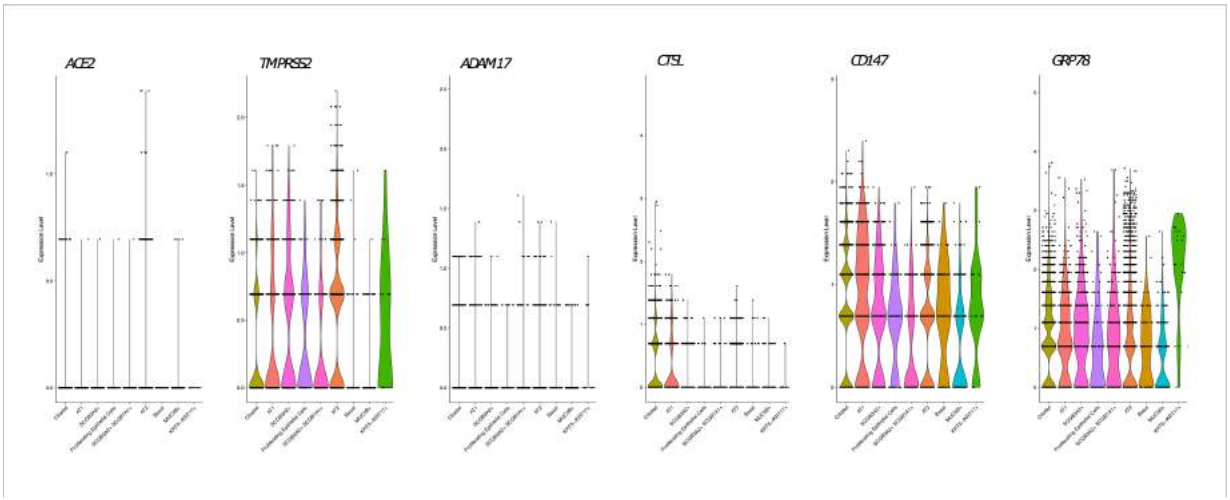


Figure 3: Violin plots for expression levels of *ACE2*, *TMPRSS2*, *ADAM17*, *CTSL*, *CD147*, and *GRP78* across lung epithelial cell populations in healthy subjects (see methods for data set reference for cell population markers).

Supplementary Figure 4

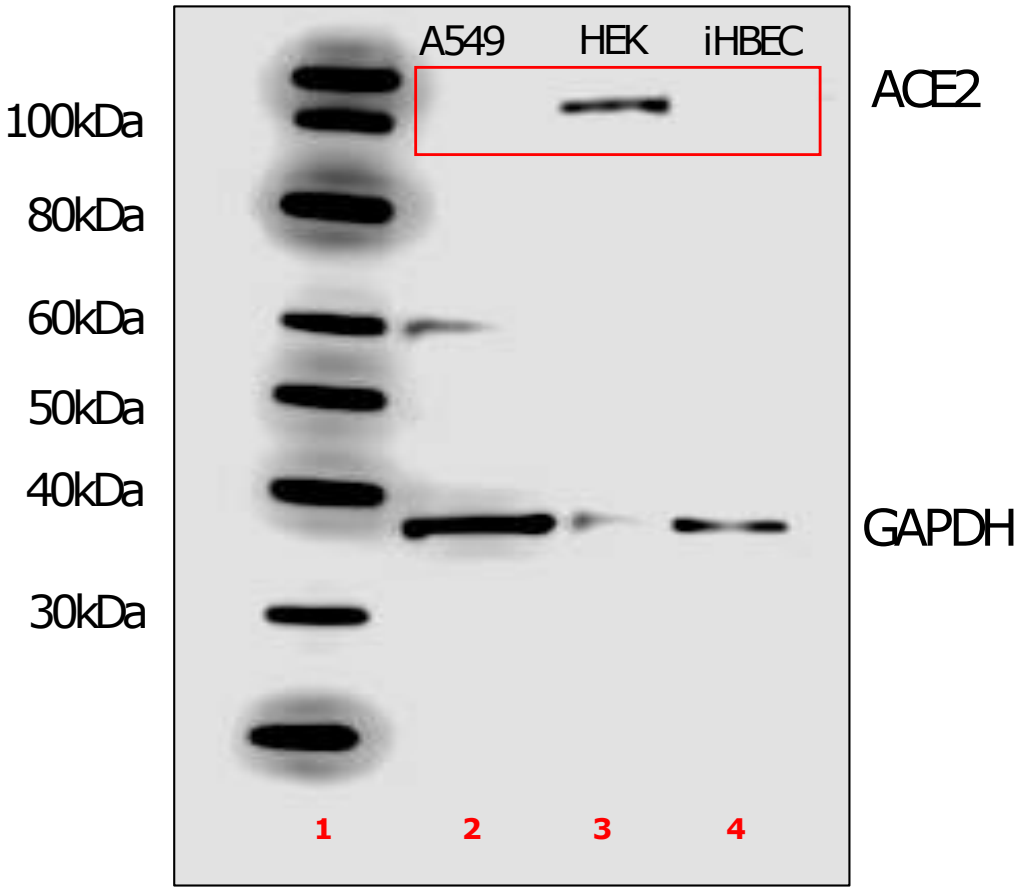


Figure 4: Lane 1: Ladder. Lane 2: A549 cell line. Lane 3: HEK293 cells. Lane 4: immortalised human bronchial epithelial cells. ACE2 has a predicted molecular weight of 110 kDa with GAPDH as a loading control. Anti-human ACE2 antibody is distinct from immunoblot in Figure 4.

Supplementary Figure 5

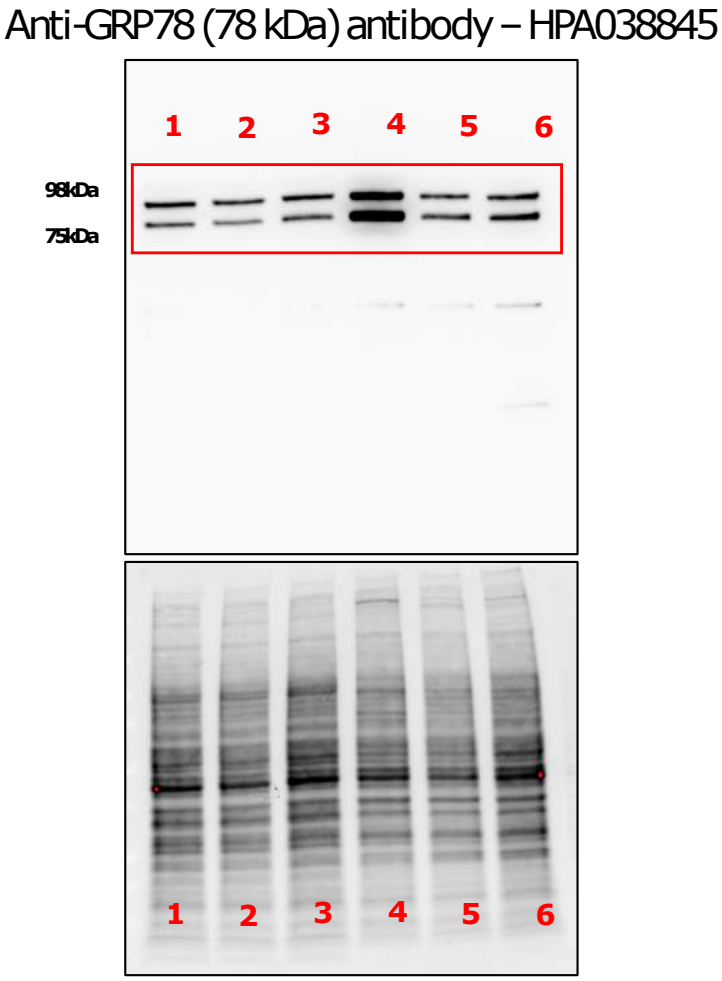


Figure 5: Lanes 1 to 3: Calu-3 cells. Lanes 4 to 6: primary human airway epithelial cells. All cells grown under submerged monolayer conditions, with $n = 3$ independent passages (Calu-3) or donor samples (primary human airway epithelial cells: non-smoker, healthy subjects). The larger band may represent GRP94 which contains the KDEL domain common with GRP78. The same samples run for this immunoblot were sampled in figure 5. Total protein loading control (bottom image) provided to demonstrate protein loaded for each sample.

Supplementary Figure 6

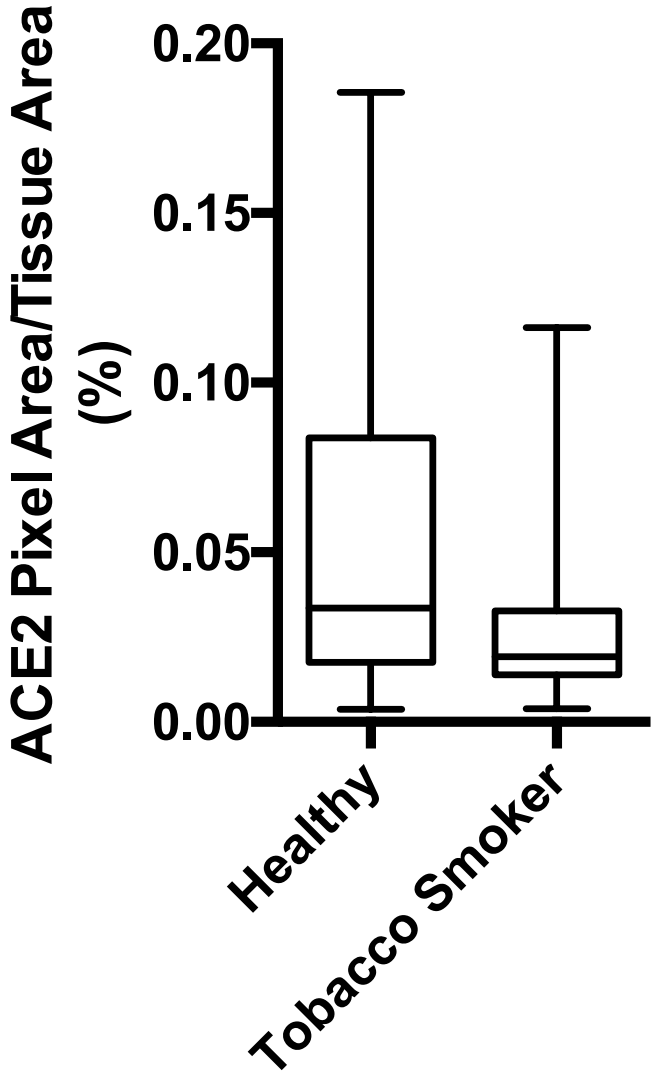


Figure 6: Positive pixels for immunohistochemical staining were expressed as a percentage of total tissue pixel count for each sample ($n = 49$). No statistical difference was observed between samples from healthy subjects and tobacco smoking subjects.

Supplementary Figure 7

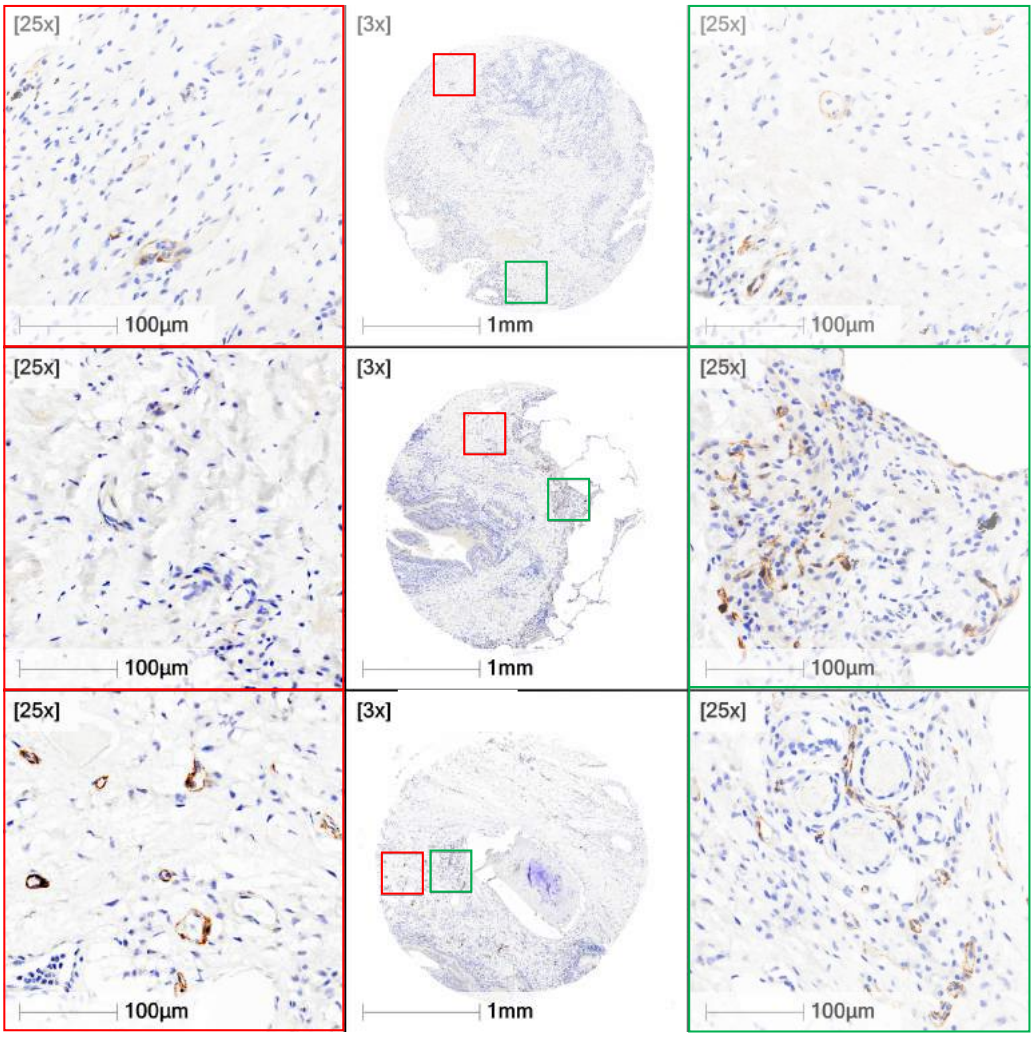


Figure 7: Representative examples ($n = 3$ donors) of positive ACE2 protein staining (rust/brown) in human lung tissue in regions distinct from those fields of view containing conducting airways. Images taken from identical slide use for Figure 5 (same staining run and conditions for image acquisition). Red and green boxes are 60× zoom of 3× magnification of entire tissue core sample.

Supplementary Figure 8

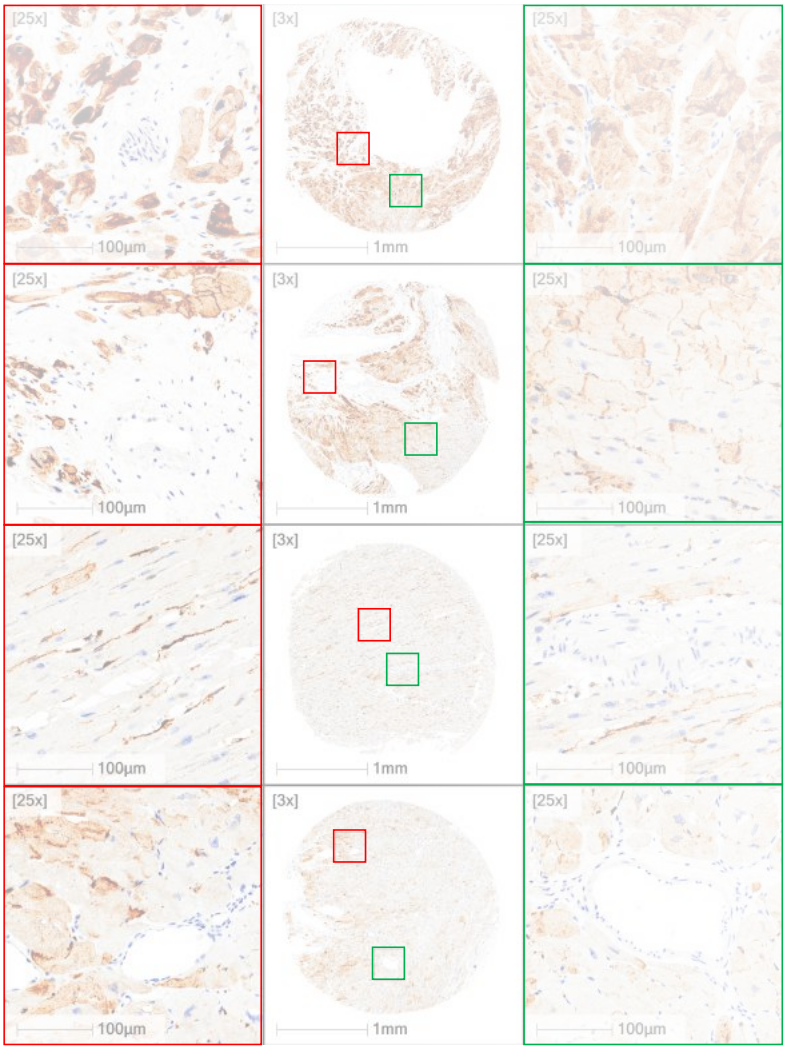


Figure 8: Representative examples ($n = 4$ donors) of positive ACE2 protein staining (rust/brown) in human heart tissue. Heart tissues stained on same staining run on Leica Bond Rx autostainer as for lung tissue. Red and green boxes are 60× zoom of 3× magnification of entire tissue core sample

Supplementary Figure 9

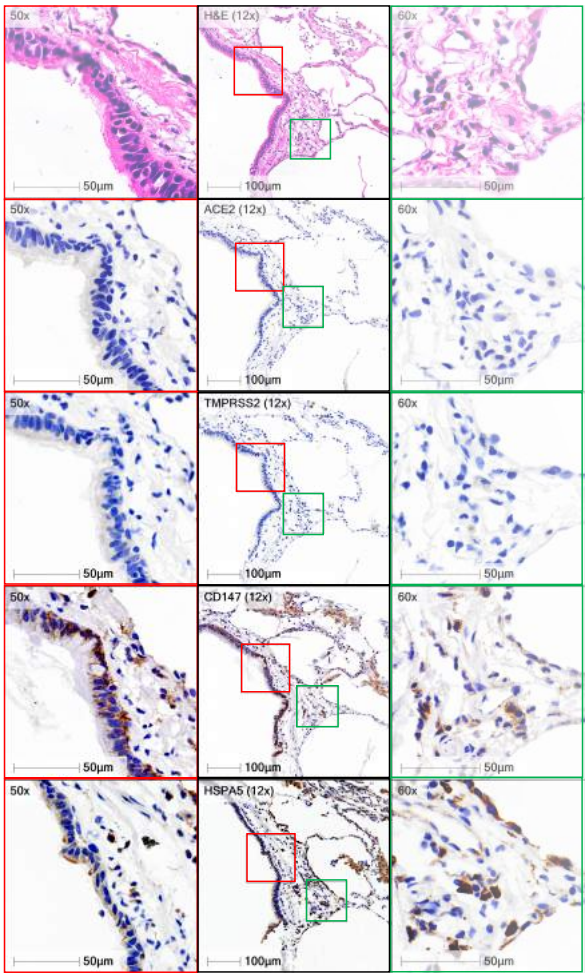
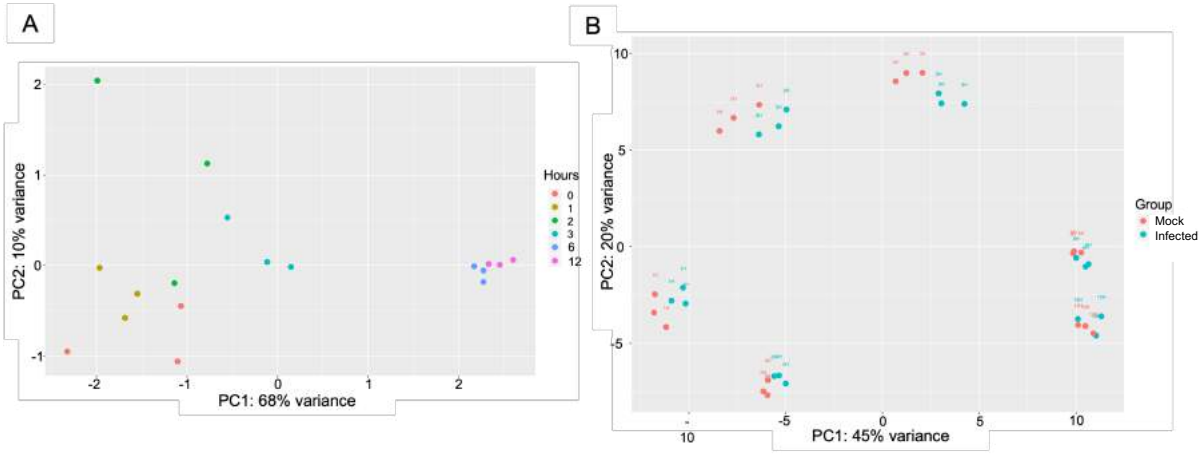


Figure 9: Black squares represent low magnification (12×) of a conducting airway with airway epithelium. Green squares correspond to high magnification regions (50×) of conducting airway epithelium that are defined in the low magnification image. Red squares correspond to high magnification regions (50×) of lung tissue away from airway lumen that are defined in the low magnification image. Row 1: hematoxylin and eosin; Row 2: ACE2; Row 3: TMPRSS2; Row 4: CD147; Row 5: GRP78/HSPA5. Positive immunohistochemical staining is rust/brown.

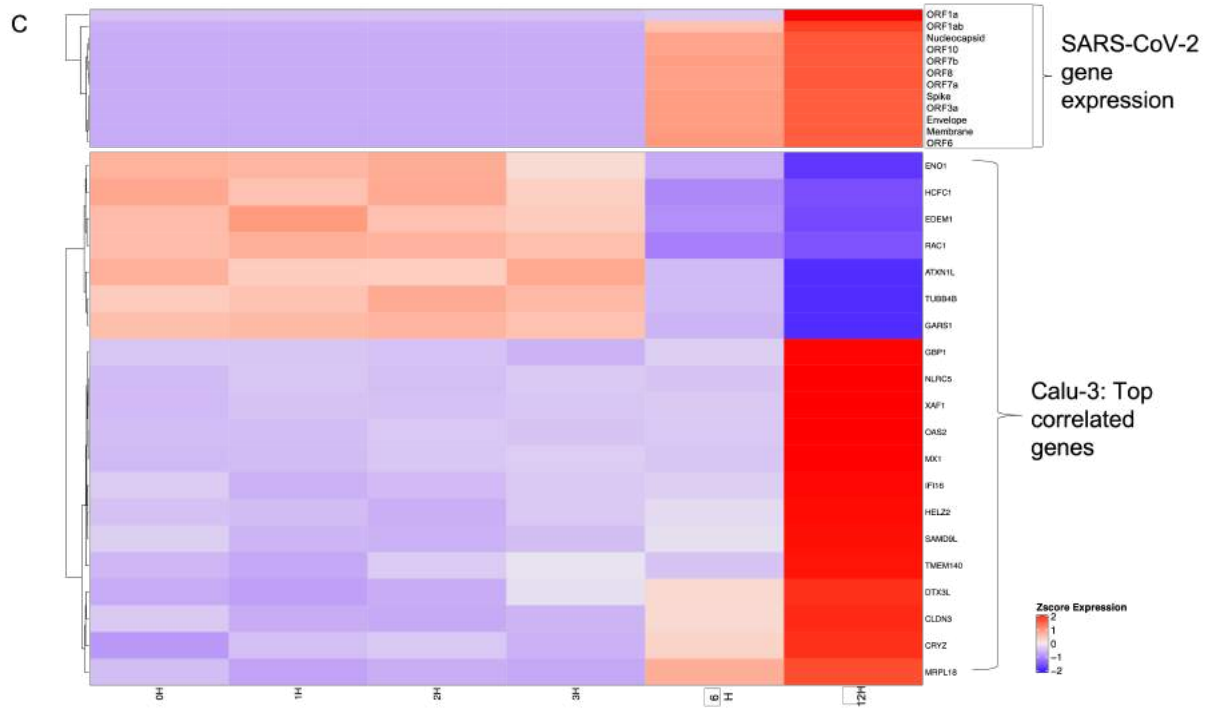
Supplementary Figure 10

Figure 10



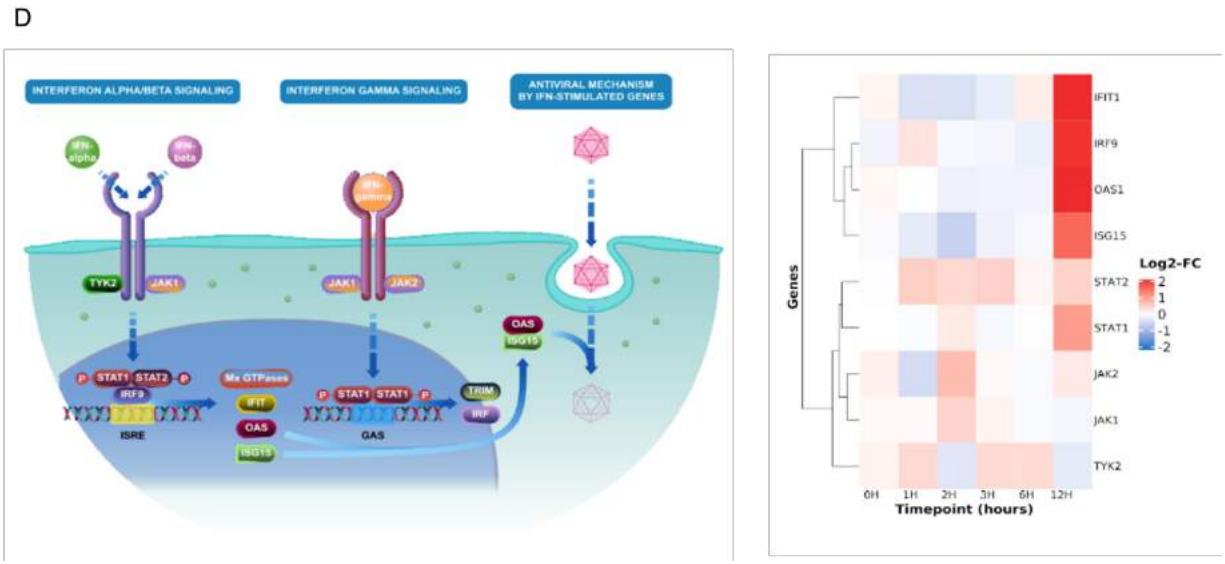
Calu-3 cells were infected at and MOI of 2. At 1 hpi, virus inoculum was replaced with cell growth media and the clock was set to 0 hours. Poly-A enriched RNA was extracted and sequenced at 0-, 1-, 2-, 3-, 6- and 12-hours post incubation (hpi). SARS-CoV-2 genome, sub-genomic RNA and transcripts were detected in infected samples. PCA clustering was performed on (A) quantified SARS-CoV-2 transcript levels in infected samples and (B) quantified and filtered host gene transcripts in both SARS-CoV-2 infected (blue) and mock infected (red) samples across time points. Axes labels indicate the proportion of between-samples variance explained by the first two principal components.

Figure 10



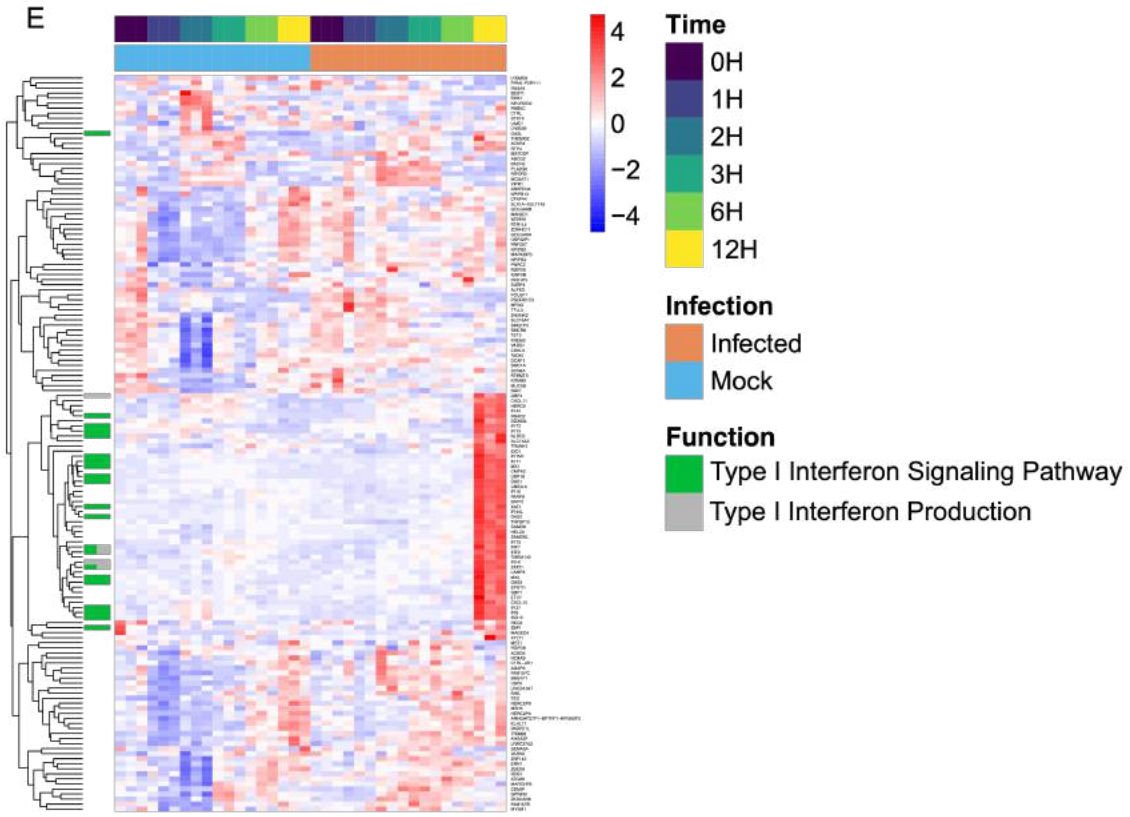
(C) Top 100 host gene expression that correlated with one or more viral transcripts over the course of infection are shown as z-score normalized expression (bottom), along with viral transcripts (top).

Figure 10



(D) Left: Pathway schematic of REACTOME cytokine signalling pathway involving interferon $\alpha/\beta/\gamma$ signalling, and antiviral response mediated by interferon stimulated genes. Right: Heat map of genes within REACTOME cytokine signalling pathway and their $\log_2$-transformed fold-change (FC) between SARS-CoV-2 infected and mock infected samples across all time points.

Figure 10



(E) Larger version of Figure 9C.

Supplementary Figure 11

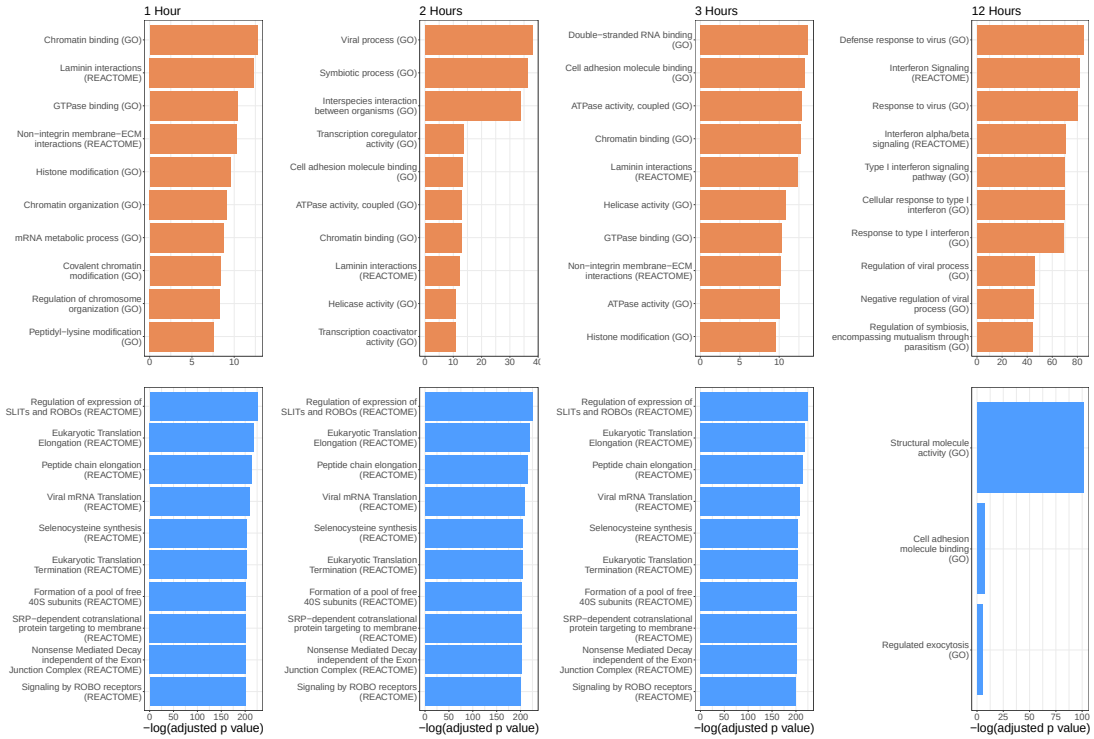
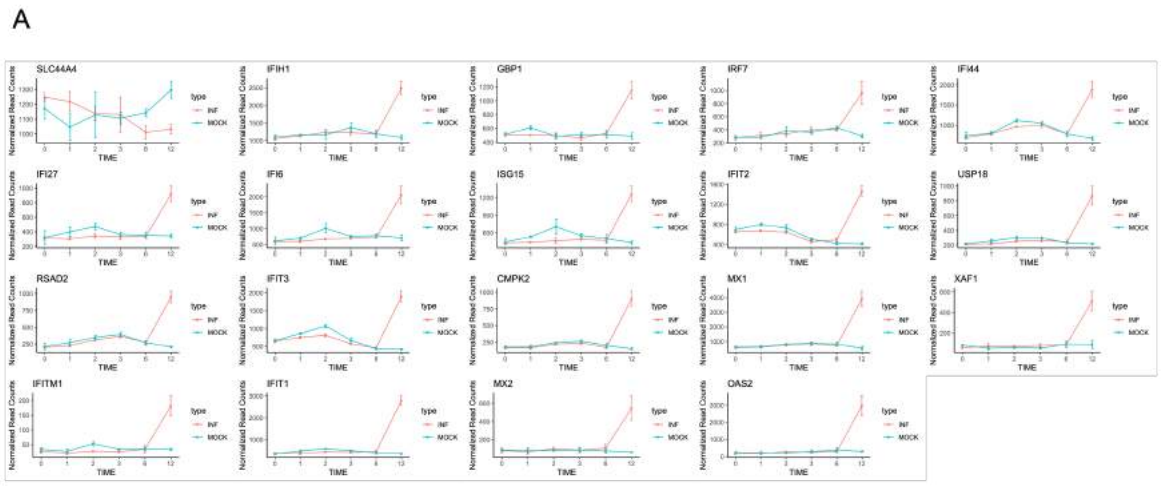


Figure 11: Top significantly (adjusted $p < 0.05$) enriched ActivePathway GO terms and REACTOME enrichments for infection versus mock at 1, 2, 3 and 12 hours post infection with SARS-CoV-2. Orange bars represent enriched terms associated with genes up-regulated in infection versus mock. Blue bars represent enriched terms associated with genes down-regulated in infection versus mock. 0 and 6 hour time points were omitted due to lack of sufficient numbers of differentially expressed genes.

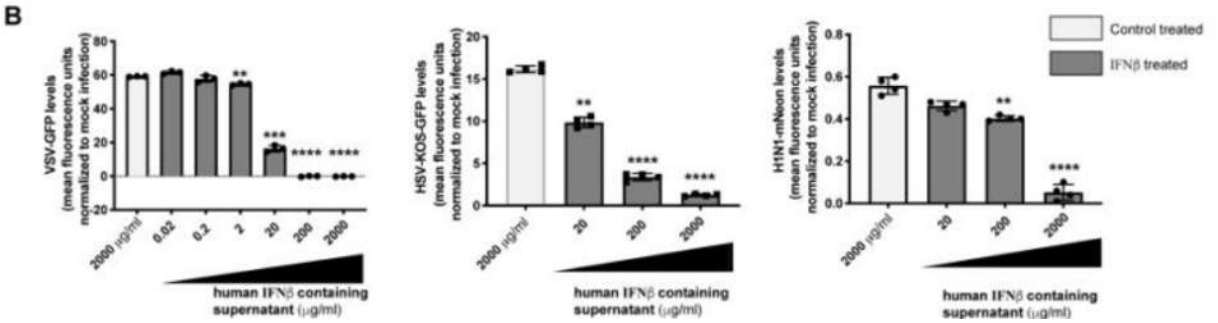
Supplementary Figure 12

Figure 12



(A) ISG expression in Calu-3 cells with significantly different levels of transcript expression between mock (blue) and SARS-CoV-2 infected (red) samples at 12 hpi are shown ($n = 3$/time point). Normalized read counts per gene, across six time-points are represented here. Time indicated is in hours. Mock, mock infected; INF, SARS-CoV-2 infected.

Figure 12



(B) Human fibroblast (THF) cells were treated with increasing concentrations of recombinant human IFNβ1 containing media or mock treated with GFP containing media (control) for 6 hours. Cells were then infected with vesicular stomatitis virus (VSV-GFP), herpes simplex virus (HSV-KOS-GFP) or H1N1 influenza virus (H1N1-mNeon). VSV and HSV were engineered to express green fluorescent protein (GFP). H1N1 expressed mNeon that is detectable in the same wavelength as GFP. Nineteen hours post incubation, GFP or mNeon levels were measured in mock infected and virus infected cells as a surrogate for virus replication. VSV-GFP ($n = 3$), HSV-KOS-GFP ($n = 4$) and H1N1-mNeon ($n = 4$) replication in THF cells treated with IFNβ1 or mock treated with control, normalized to mock infection is shown above. Data are represented as mean +/- SD, $n = 3$, 4 or 6, p: ** < 0.01, *** < 0.001 and **** < 0.0001 (Student's t test). GFP and mNeon expression is represented after normalization with mock infected cells. hpi, hours post incubation.

Supplementary Figure 13

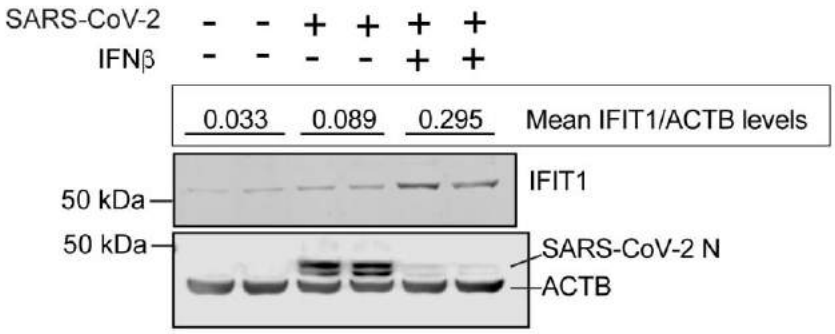


Figure 13: IFIT1, SARS-CoV-2 N and ACTB protein expression in Calu-3 cells that were infected with SARS-CoV-2 or mock infected for 1 hour, followed by control or IFNβ treatment for 72 hours ($n = 2$). Data are represented as mean +/- SD, $n = 2$. IFIT1 protein expression levels are expressed as ratios of IFIT1/ACTB levels. Blots were quantified using Image Studio (Li-COR) ($n = 2$).

Supplementary Table 10

Table 10: Metadata and references for curated GEO deposited data sets of primary human airway epithelial cells

GSE Accession	Affymetrix Chip	Sample Annotations (N=504)	Citation
GSE4302	HG-U133_Plus_2	15 current smokers 11 never smokers 2 unknown	Woodruff et al., 2007; 2009
GSE37147	HuGene-1_0-st-v1	82 former smokers (49 male, 33 female) 69 current smokers (34 male, 35 female) 8 unknown	Steiling et al., 2013
GSE67472	HG-U133_Plus_2	23 male 20 female	Christenson et al., 2015
GSE108134	HG-U133_Plus_2	176 current smokers 98 never smokers	O'Beirne et al., 2018

Supplementary Table 11

Table 11: Statistics for impact of smoking status relative to never smokers on gene expression of SARS-CoV-2 candidate receptors

	ACE2	TMPRSS2	ADAM17	CTSL	CD147	GRP78	CDH1
Current vs Former	***	***	*	***	***		**
Current vs Never	***	**		***	***	*	***
Former vs Never	*	**			***		***

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Results for each individual gene obtained via ANOVA followed by Tukey's HSD *post-hoc test* using the stats package in R (version 3.2.1).

Supplementary Table 12

Table 12: Mean raw read counts for SARS-CoV-2 transcripts. INF, SARS-CoV-2 infected; H, hours post incubation; SD, standard deviation.

Mean INF 0H	SD INF 0H	Mean INF 1H	SD INF 1H	Mean INF 2H	SD INF 2H	Mean INF 3H	SD INF 3H	Mean INF 6H	SD INF 6H	Mean INF 12H	SD INF 12H	SARS-CoV-2 gene	Transcript
257.67	38.59	285.33	56.13	243.67	39.25	278.00	23.00	12173.3	3006.93	25827.33	2054.93	ORF1ab	lcl|NC_045512.2_cds_YP_009724389.1_1
1.00	1.73	0.00	0.00	0.33	0.58	0.00	0.00	33.67	6.81	1061.00	468.03	ORF1a	lcl|NC_045512.2_cds_YP_009725295.1_2
500.67	94.52	491.33	86.19	378.00	61.39	521.67	49.69	19232.3	3952.46	26903.33	3860.82	spike	lcl|NC_045512.2_cds_YP_009724390.1_3
173.67	24.99	172.33	43.68	127.33	17.16	203.33	26.50	9995.00	1736.00	13976.33	2233.55	ORF3a	lcl|NC_045512.2_cds_YP_009724391.1_4
43.67	5.51	44.67	13.65	39.00	2.65	63.00	11.53	2903.33	485.15	4086.33	627.70	envelope	lcl|NC_045512.2_cds_YP_009724392.1_5
199.67	27.02	196.00	37.32	162.33	28.87	298.67	19.60	22344.3	3354.18	31200.33	4915.23	membrane	lcl|NC_045512.2_cds_YP_009724393.1_6
34.67	2.08	32.33	10.50	25.00	7.81	45.33	1.53	3508.00	509.12	4704.67	886.56	ORF6	lcl|NC_045512.2_cds_YP_009724394.1_7
107.33	19.50	102.33	23.35	94.00	22.61	173.67	34.00	14834.0	2357.53	21920.67	3441.71	ORF7a	lcl|NC_045512.2_cds_YP_009724395.1_8
10.33	2.52	11.67	2.31	15.33	1.53	20.67	1.15	1516.33	241.00	2191.33	526.17	ORF7b	lcl|NC_045512.2_cds_YP_009725318.1_9
109.33	22.19	107.00	27.22	98.00	21.70	189.00	14.00	14651.3	2136.80	21518.67	3992.04	ORF8	lcl|NC_045512.2_cds_YP_009724396.1_10
1251.00	230.97	1157.33	247.52	1067.67	144.58	2945.67	402.61	258553	34843.96	393221.67	62159.07	nucleocapsid	lcl|NC_045512.2_cds_YP_009724397.2_11
112.33	27.57	97.00	22.52	94.67	10.69	250.00	19.00	18385.3	2239.71	27679.00	5406.01	ORF10	lcl|NC_045512.2_cds_YP_009725255.1_12

Supplementary Table 13

Table 13: Mean normalized read counts for differentially expressed IFN and ISG transcripts. H, hours post incubation; INF, SARS-CoV-2 infected; MOCK, mock infected; IFN, interferon; ISG, interferon stimulated genes.

		0H INF (N=3)	0H MOCK (N=3)	1H INF (N=3)	1H MOCK (N=3)	2H INF (N=3)	2H MOCK (N=3)	3H INF (N=3)	3H MOCK (N=3)	6H INF (N=3)	6H MOCK (N=3)	12H INF (N=3)	12H MOCK (N=3)
IFNs	IFNB1	1.35	0.00	1.21	0.41	1.48	0.97	0.57	1.93	6.40	0.30	21.23	0.89
	IFNL1	3.49	3.45	2.20	4.80	4.93	5.46	3.17	1.90	7.00	2.66	15.07	0.73
	IFNL2	0.00	0.00	0.36	0.96	4.11	0.35	0.28	0.00	4.66	0.00	8.61	0.00
	IFNL3	0.35	0.00	0.58	0.44	2.38	0.31	0.88	0.00	3.02	0.00	8.46	0.00
ISGs	IFIT1	388.42	358.77	370.80	487.33	447.59	590.32	425.31	498.05	463.17	408.65	2790.57	367.50
	IRF7	278.50	283.73	320.43	284.00	339.99	383.89	399.07	363.67	399.93	432.29	966.54	305.31
	OAS2	172.67	236.24	178.18	222.85	287.61	208.20	252.85	296.10	292.36	378.90	2979.22	303.60
	MX1	588.48	620.75	624.79	647.52	758.95	800.13	839.47	867.29	728.29	811.68	3922.41	546.94
	RSAD2	204.76	216.53	228.73	272.67	313.84	348.31	365.12	393.68	274.53	269.56	948.75	210.54
	SLC44A4	1247.82	1171.72	1218.77	1046.17	1138.09	1128.19	1129.60	1106.06	1010.30	1142.19	1032.09	1298.09
	IFIH1	1052.81	1100.39	1134.78	1163.76	1235.36	1164.31	1223.66	1371.55	1189.70	1191.00	2492.69	1087.88
	GBP1	506.79	512.73	503.57	608.29	496.74	485.28	458.14	509.15	530.04	509.53	1151.35	488.92
	IFI44	689.16	741.40	789.19	803.61	963.68	1113.99	997.06	1052.67	785.42	782.39	1889.54	671.51
	IFI27	311.49	318.74	302.63	399.59	343.37	472.30	328.28	361.48	333.63	351.85	921.55	342.54
	IFI6	592.82	612.04	599.90	697.80	673.06	1010.20	692.26	752.25	729.19	775.17	2066.30	709.85
	ISG15	430.95	447.57	443.60	533.02	465.88	704.43	490.49	554.07	473.88	502.97	1260.48	435.91
	IFIT2	657.23	698.02	676.46	795.49	645.57	732.08	455.75	504.29	493.48	422.04	1465.16	413.27
	USP18	212.27	217.53	218.01	257.03	253.55	301.50	266.17	297.44	243.57	232.66	873.18	218.27
	IFIT3	648.15	656.89	747.61	858.17	810.13	1069.67	567.26	668.13	458.25	428.90	1900.07	420.64
	CMPK2	163.89	179.41	169.11	182.05	219.35	244.03	235.97	265.54	172.78	201.60	906.22	153.23
	XAF1	58.53	82.76	73.40	53.61	69.79	60.14	79.67	55.09	86.30	91.97	513.01	90.51
	IFITM1	27.68	34.25	21.94	27.89	28.53	53.49	26.88	34.91	34.59	35.75	182.01	34.33
	MX2	82.11	87.24	69.22	81.96	100.75	83.43	87.84	87.48	108.05	78.88	547.98	64.92

Supplementary Table 14

Table 14: Pathway enrichment analysis. Significance was determined after FDR correction. H, hours post incubation; 0, non-significant; 1, significant.

Term ID	Term Name	Adjusted *p* value	1H	2H	3H	12H
GO:0000976	transcription regulatory region sequence-specific DNA binding	0.004824255	0	1	0	0
GO:0001067	regulatory region nucleic acid binding	0.004203707	0	1	0	0
GO:0001816	cytokine production	0.005529472	0	0	0	1
GO:0001817	regulation of cytokine production	0.001829233	0	0	0	1
GO:0002230	positive regulation of defense response to virus by host	0.002197834	0	0	0	1
GO:0002831	regulation of response to biotic stimulus	8.60E-08	0	0	0	1
GO:0002833	positive regulation of response to biotic stimulus	0.008687053	0	0	0	1
GO:0003690	double-stranded DNA binding	0.000112873	0	1	0	0
GO:0003712	transcription coregulator activity	1.30E-06	0	1	0	0
GO:0003713	transcription coactivator activity	2.39E-05	0	1	0	0
GO:0005178	integrin binding	0.013874905	0	0	1	0
GO:0008270	zinc ion binding	0.000103938	0	1	0	0
GO:0009615	response to virus	1.39E-35	0	0	0	1
GO:0010810	regulation of cell-substrate adhesion	0.008350323	0	1	0	0
GO:0016482	cytosolic transport	0.011086056	0	1	0	0
GO:0019058	viral life cycle	3.92E-11	0	0	0	1
GO:0019079	viral genome replication	3.87E-15	0	0	0	1
GO:0019221	cytokine-mediated signaling pathway	8.45E-16	0	0	0	1
GO:0019900	kinase binding	0.003539788	0	1	0	0
GO:0019901	protein kinase binding	0.012867428	0	1	0	0
GO:0030099	myeloid cell differentiation	0.011382292	0	1	0	0
GO:0031347	regulation of defense response	2.16E-05	0	0	0	1
GO:0031589	cell-substrate adhesion	0.002867293	0	1	0	0
GO:0032020	ISG15-protein conjugation	0.008627708	0	0	0	1
GO:0032069	regulation of nuclease activity	1.26E-06	0	0	0	1
GO:0032479	regulation of type I interferon production	4.92E-06	0	0	0	1
GO:0032480	negative regulation of type I interferon production	0.005210998	0	0	0	1
GO:0032481	positive regulation of type I interferon production	0.00531473	0	0	0	1
GO:0032606	type I interferon production	6.14E-06	0	0	0	1
GO:0032607	interferon-alpha production	0.005237546	0	0	0	1
GO:0032647	regulation of interferon-alpha production	0.00400414	0	0	0	1
GO:0032727	positive regulation of interferon-alpha production	0.001567461	0	0	0	1
GO:0034340	response to type I interferon	9.21E-31	0	0	0	1
GO:0034341	response to interferon-gamma	1.44E-10	0	0	0	1

GO:0034504	protein localization to nucleus	0.00295333	0	1	0	0
GO:0035455	response to interferon-alpha	3.29E-10	0	0	0	1
GO:0035456	response to interferon-beta	2.08E-07	0	0	0	1
GO:0042393	histone binding	0.00298728 5	0	1	0	0
GO:0043900	regulation of multi-organism process	2.03E-17	0	0	0	1
GO:0043901	negative regulation of multi-organism process	3.86E-17	0	0	0	1
GO:0043902	positive regulation of multi-organism process	0.00827448 4	0	0	0	1
GO:0043903	regulation of symbiosis encompassing mutualism through parasitism	6.66E-20	0	0	0	1
GO:0044212	transcription regulatory region DNA binding	0.00404741 6	0	1	0	0
GO:0045069	regulation of viral genome replication	1.01E-16	0	0	0	1
GO:0045071	negative regulation of viral genome replication	3.61E-17	0	0	0	1
GO:0045088	regulation of innate immune response	5.98E-06	0	0	0	1
GO:0045089	positive regulation of innate immune response	0.00597980 2	0	0	0	1
GO:0046596	regulation of viral entry into host cell	0.04802533 7	0	0	0	1
GO:0048525	negative regulation of viral process	3.26E-20	0	0	0	1
GO:0050657	nucleic acid transport	0.04848561 5	0	0	1	0
GO:0050658	RNA transport	0.04848561 5	0	0	1	0
GO:0050688	regulation of defense response to virus	0.00216321 6	0	0	0	1
GO:0050691	regulation of defense response to virus by host	0.00954189 2	0	0	0	1
GO:0050792	regulation of viral process	1.31E-20	0	0	0	1
GO:0051056	regulation of small GTPase mediated signal transduction	0.02604849 5	0	1	0	0
GO:0051607	defense response to virus	1.25E-37	0	0	0	1
GO:0060333	interferon-gamma-mediated signaling pathway	1.36E-13	0	0	0	1
GO:0060337	type I interferon signaling pathway	3.69E-31	0	0	0	1
GO:0060700	regulation of ribonuclease activity	6.89E-07	0	0	0	1
GO:0060759	regulation of response to cytokine stimulus	0.00074017 3	0	0	0	1
GO:0060760	positive regulation of response to cytokine stimulus	0.00710556 4	0	0	0	1
GO:0061629	RNA polymerase II-specific DNA-binding transcription factor binding	0.01112665 6	0	1	0	0
GO:0070566	adenylyltransferase activity	0.00654540 2	0	0	0	1
GO:0071346	cellular response to interferon-gamma	1.05E-09	0	0	0	1
GO:0071357	cellular response to type I interferon	3.69E-31	0	0	0	1
GO:0098586	cellular response to virus	0.0037813	0	0	0	1
GO:1903900	regulation of viral life cycle	1.50E-18	0	0	0	1
GO:1903901	negative regulation of viral life cycle	1.15E-18	0	0	0	1
GO:1990837	sequence-specific double-stranded DNA binding	0.00294552 6	0	1	0	0
GO:2001251	negative regulation of chromosome organization	0.03997967 2	0	1	0	0
REAC:R-HSA-1169408	ISG15 antiviral mechanism	5.61E-12	0	0	0	1

REAC:R-HSA-1169410	Antiviral mechanism by IFN-stimulated genes	5.77E-19	0	0	0	1
REAC:R-HSA-1280215	Cytokine Signaling in Immune system	1.52E-19	0	0	0	1
REAC:R-HSA-168928	DDX58/IFIH1-mediated induction of interferon-alpha/beta	0.001851135	0	0	0	1
REAC:R-HSA-2990846	SUMOylation	0.000289223	0	1	0	0
REAC:R-HSA-3108214	SUMOylation of DNA damage response and repair proteins	0.023406467	0	1	0	0
REAC:R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	0.000850049	0	1	0	0
REAC:R-HSA-3247509	Chromatin modifying enzymes	0.016088428	0	1	0	0
REAC:R-HSA-4839726	Chromatin organization	0.016088428	0	1	0	0
REAC:R-HSA-6806834	Signaling by MET	2.89E-05	0	1	0	0
REAC:R-HSA-877300	Interferon gamma signaling	2.97E-09	0	0	0	1
REAC:R-HSA-8874081	MET activates PTK2 signaling	0.000994797	1	0	0	0
REAC:R-HSA-8934593	Regulation of RUNX1 Expression and Activity	0.000745328	0	1	0	0
REAC:R-HSA-8983711	OAS antiviral response	3.29E-08	0	0	0	1
REAC:R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	0.017643755	0	1	0	0
REAC:R-HSA-909733	Interferon alpha/beta signaling	2.97E-31	0	0	0	1
REAC:R-HSA-913531	Interferon Signaling	4.75E-36	0	0	0	1
REAC:R-HSA-918233	TRAF3-dependent IRF activation pathway	0.000139967	0	0	0	1
REAC:R-HSA-933541	TRAF6 mediated IRF7 activation	0.018776243	0	0	0	1
REAC:R-HSA-936440	Negative regulators of DDX58/IFIH1 signaling	0.000931238	0	0	0	1

Supplementary Table 15

Table 15: COVID-19 patient serum sample history and sera from healthy controls.

Sample ID	Sample Type	Age (years)	Sex	Charlson co-morbidity score (weighed sum of co-morbidity groups*)	Immunosuppressed	Admitted to hospital	Admission date	Required oxygen	ICU	ICU admission date	Outcome	Death/ discharge date	Disease severity
COVIFN1	serum - acute	59	M	0	N	Y	2020-04-11	Y	N	NA	Discharged alive	2020-04-19	Moderate
COVIFN2	serum - acute	86	M	8	N	Y	2020-04-17	N	N	NA	Discharged alive	2020-04-29	Moderate
COVIFN3	serum - acute	96	F	2	N	Y	2020-04-26	Y	N	NA	Discharged alive	2020-05-28	Moderate
COVIFN4	serum - acute	57	M	0	N	Y	2020-04-23	Y	N	NA	Discharged alive	2020-05-06	Moderate
COVIFN5	serum - acute	54	M	1	N	Y	2020-04-28	N	N	NA	Discharged alive	2020-05-07	Moderate
COVIFN6	serum - acute	62	F	6	Y	Y	2020-04-26	N	N	NA	Discharged alive	2020-05-01	Moderate
COVIFN7	serum - acute	83	F	3	N	Y	2020-04-28	Y	N	NA	Discharged alive	2020-05-22	Moderate
COVIFN8	serum - acute	85	M	1	N	Y	2020-05-02	N	N	NA	Discharged alive	2020-05-05	Moderate
COVIFN9	serum - acute	23	M	1	N	Y	2020-04-29	Y	N	NA	Discharged alive	2020-05-11	Moderate
COVIFN10	serum - acute	58	F	4	Y	Y	2020-05-03	N	N	NA	Discharged alive	2020-05-07	Moderate
COVIFN11	serum - acute	80	M	0	N	Y	2020-04-02	Y	Y	2020-04-03	Discharged alive	2020-06-12	Severe
COVIFN12	serum - acute	69	F	0	N	Y	2020-04-08	Y	Y	2020-04-10	Died	2020-04-15	Severe
COVIFN13	serum - acute	65	M	0	N	Y	2020-04-07	N	Y	2020-04-17	Discharged alive	2020-04-28	Severe
COVIFN14	serum - acute	73	M	0	N	Y	2020-04-06	Y	N	NA	Discharged alive	2020-04-23	Severe
COVIFN15	serum - acute	81	M	2	N	Y	2020-04-16	Y	Y	2020-04-16	Died	2020-04-29	Severe
COVIFN16	serum - acute	72	M	1	N	Y	2020-04-18	Y	Y	2020-04-18	Died	2020-05-27	Severe
COVIFN17	serum - acute	49	M	0	N	Y	2020-04-24	Y	Y	2020-04-26	Discharged alive	2020-05-11	Severe
COVIFN18	serum - acute	75	M	8	N	Y	2020-04-26	Y	Y	2020-04-27	Died	2020-05-19	Severe
COVIFN19	serum - acute	94	F	3	N	Y	2020-04-08	Y	N	NA	Died	2020-05-16	Severe
COVIFN20	serum - acute	80	F	8	N	Y	2020-04-14	Y	N	NA	Died	2020-04-29	Severe
OM1	NA	59	M	NA	NA	NA	NA	NA	NA	NA	NA	NA	Healthy
OM8035	NA	26	M	NA	NA	NA	NA	NA	NA	NA	NA	NA	Healthy
OM908	NA	30	F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Healthy
OM920	NA	58	M	NA	NA	NA	NA	NA	NA	NA	NA	NA	Healthy
OM921	NA	36	F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Healthy

***Charlson co-morbidity groups: 1: Myocardial infarction, 2: Congestive heart failure, 3: Peripheral vascular disease, 4: Cerebrovascular disease, 5: Dementia, 6: COPD, 7: Rheumotological disease, 8: Peptic ulcer disease, 9: Mild liver disease, 10: Diabetes without complications, 11: Diabetes with complications, 12: Paraplegia and hemiplegia, 13: Renal disease, 14: Malignancy, including leukemia, 15: Moderate or severe liver disease, 16: Metastatic cancer, 17: HIV/AIDS**

Supplementary Table 16

Table 16: Cytokine levels in healthy individuals and COVID-19 patient serum samples.

Sample ID/Cytokine level (pg/ml)	COVIFN1	COVIFN2	COVIFN3	COVIFN4	COVIFN5	COVIFN6	COVIFN7	COVIFN8	COVIFN9	COVIFN10	COVIFN11	COVIFN12
sCD40L	1.24	11.05	784.48	2598.72	4394.41	1.24	1629.28	1500.89	4096.10	1441.99	1362.29	8.15
EGF	137.74	142.39	116.09	623.86	293.83	244.91	128.46	24.00	347.18	125.25	120.85	66.21
Eotaxin	6.81	9.76	15.32	55.68	19.80	2.81	26.02	27.77	16.54	43.88	36.36	4.95
FGF-2	31.41	35.81	37.87	48.89	42.71	31.41	77.97	36.85	209.35	41.77	36.85	26.52
FLT-3L	25.20	18.33	7.16	8.10	11.44	13.55	7.54	20.43	18.87	44.95	4.00	24.72
Fractalkine	17.13	35.95	25.13	83.63	13.96	10.31	1334.65	25.13	145.58	43.41	336.72	46.88
G-CSF	OOR <	OOR <	OOR <	OOR <	OOR <	OOR <	11.74	OOR <	17.24	277.59	41.61	OOR <
GM-CSF	7.04	7.86	11.89	6.22	1.79	3.17	32.11	5.96	24.92	18.69	16.79	14.31
GROα	10.12	8.84	7.15	17.41	11.50	6.65	12.54	8.25	86.48	19.96	20.22	44.06
IFN-α2	OOR <	OOR <	47.64	66.74	7.36	OOR <	1864.77	7.36	44.39	OOR <	23.75	OOR <
IFNγ	OOR <	1.02	1.79	0.28	0.07	OOR <	22.35	0.70	2.92	4.29	2.15	OOR <
IL-1α	3.52	17.68	0.03	10.84	1.16	OOR <	6.85	1.74	29.05	1.16	5.33	2.04
IL-1β	4.29	8.94	1.04	12.45	2.81	2.38	9.33	4.91	24.47	2.81	1.50	3.45
IL-1RA	1.75	1.75	34.98	22.97	5.99	0.72	17.76	226.55	9.76	243.02	22.47	0.30
IL-2	OOR <	OOR <	OOR <	0.55	OOR <	OOR <	3.96	OOR <	1.34	0.45	1.22	OOR <
IL-3	OOR <	OOR <	OOR <	OOR <	OOR <	0.03	185.31	OOR <	0.03	0.03	OOR <	OOR <
IL-4	1.90	0.60	0.67	0.44	0.60	OOR <	0.21	0.23	2.49	0.25	0.20	0.49
IL-5	7.43	9.64	4.29	2.96	4.82	0.26	2.48	8.93	4.17	4.09	54.76	33.72
IL-6	13.87	7.04	22.47	7.19	3.50	9.56	9.13	2.15	4.84	57.07	25.96	158.27
IL-7	5.40	1.46	1.04	15.61	12.91	3.52	104.40	5.56	9.76	7.96	3.52	14.03
IL-8	9.31	7.99	6.85	22.33	11.85	2.06	14.04	5.36	10.76	15.26	21.77	8.71
IL-9	OOR <	5.23	OOR <	OOR <	4.51	OOR <	6.02	OOR <	17.85	4.04	3.17	4.94
IL-10	OOR <	OOR <	11.70	3.72	0.64	OOR <	13.54	0.80	6.60	338.81	18.13	OOR <
IL-12p40	1.63	12.02	53.59	18.91	OOR <	OOR <	91.32	25.65	66.30	7.98	644.91	OOR <
IL-12p70	0.05	OOR <	0.19	OOR <	OOR <	0.34	0.19	0.49	5.27	0.49	0.05	OOR <
IL-13	OOR <	6.03	16.52	35.11	45.40	OOR <	6.03	35.74	135.56	OOR <	151.34	3.84
IL-15	9.11	8.90	13.60	7.30	8.36	23.59	11.57	7.30	21.44	83.09	7.51	22.73
IL-17A	OOR <	3.54	1.43	9.30	1.87	1.87	2.51	1.65	8.77	0.02	1.08	OOR <
IL-17E/IL-25	150.37	445.50	39.60	OOR <	OOR <	OOR <	244.48	OOR <	583.89	OOR <	OOR <	87.52
IL-17F	OOR <	OOR <	2.68	1.54	1.06	OOR <	5.90	2.91	154.23	0.82	OOR <	OOR <
IL-18	1.51	10.63	20.75	23.25	17.64	11.38	24.52	10.18	35.10	56.13	28.32	69.33
IL-22	OOR <	OOR <	OOR <	OOR <	OOR <	OOR <	OOR <	OOR <	107.30	OOR <	108.47	OOR <
IL-27	OOR <	OOR <	1174.89	1342.14	616.07	OOR <	5779.33	2252.24	2404.96	2396.84	1290.93	53.99
IP-10	101.29	171.41	408.11	306.80	337.00	414.99	3692.55	2872.70	88.11	6241.52	485.68	3403.51
MCP-1	287.95	156.83	215.68	274.98	153.75	184.71	191.73	217.00	33.06	594.23	276.08	527.77
MCP-3	5.05	4.46	17.04	13.20	8.08	OOR <	7.90	7.34	25.71	10.03	25.27	22.17
M-CSF	87.77	82.16	102.43	189.92	86.65	421.41	196.93	64.34	141.25	505.58	100.17	262.88
MDC	173.17	234.10	262.67	364.53	327.27	15.60	112.12	518.15	466.85	89.15	660.76	215.22
MIG/CXCL9	1466.74	4016.77	2018.24	1067.21	1816.00	654.79	2752.32	8500.06	1134.79	5079.75	5120.21	5694.96
MIP-1α	20.80	14.42	14.97	32.09	17.04	5.53	24.21	11.50	39.29	18.02	14.42	18.02
MIP-1β	26.59	28.34	24.01	37.83	31.00	9.91	19.70	44.87	46.08	48.28	73.21	34.57
PDGF-AA	690.76	1475.89	1061.65	977.84	2208.47	283.99	941.93	1429.40	2550.13	1002.92	1000.69	1105.14
PDGF-AB/BB	17241.96	17026.36	12565.55	13503.22	18036.00	10027.45	12919.49	13950.58	13653.00	14280.43	11440.21	15560.94
RANTES	838.06	856.79	779.87	790.28	647.88	805.26	1026.05	702.34	658.19	858.07	677.35	829.54
TGFα	1.63	2.43	2.69	2.99	3.38	3.13	3.91	3.28	77.54	7.64	12.83	5.47
TNFα	4.80	11.80	65.20	52.38	33.86	35.09	59.25	79.93	122.76	122.76	138.16	56.57
TNFβ	2.20	2.05	3.04	14.47	10.09	OOR <	5.89	8.13	16.23	1.78	26.15	4.68
VEGF-A	12.49	122.01	54.27	271.90	335.94	8.53	118.23	299.64	108.40	103.15	279.12	262.42

Sample ID/Cytokine level (pg/ml)	COVIFN13	COVIFN14	COVIFN15	COVIFN16	COVIFN17	COVIFN18	COVIFN19	COVIFN20	OM1	OM8035	OM908	OM920	OM921
sCD40L	1549.90	2350.61	1002.59	226.82	55.64	1094.25	2476.20	13.78	255.18	117.26	35.09	74.59	258.68
EGF	19.33	140.45	92.68	68.63	726.00	113.98	244.36	71.11	18.11	15.19	OOR <	22.51	25.62
Eotaxin	27.77	73.74	22.87	64.36	8.01	44.64	29.03	27.56	29.70	36.89	17.38	55.40	25.15
FGF-2	110.76	50.56	38.87	84.58	81.02	103.20	47.18	42.71	287.46	74.85	26.52	84.58	37.87
FLT-3L	4.97	16.53	35.05	23.35	31.15	24.32	25.51	20.07	16.62	3.21	3.95	11.62	3.80
Fractalkine	120.95	35.95	116.92	77.37	79.91	215.50	119.95	46.88	122.94	116.92	118.94	110.72	88.46
G-CSF	47.51	132.14	52.98	0.99	7.29	103.19	51.42	OOR <	45.92	2.81	OOR <	18.71	10.16
GM-CSF	42.84	19.01	30.37	20.30	4.91	23.59	26.94	10.42	16.79	5.69	0.57	18.06	12.79
GROα	31.61	28.79	21.91	15.86	74.17	18.16	24.20	79.24	11.12	2.10	2.01	3.39	18.45
IFN-α2	68.09	35.73	30.02	28.01	142.39	142.39	98.86	13.84	176.90	39.30	81.06	46.03	31.97
IFNγ	2.61	1.17	0.94	0.55	2.79	1.77	0.81	0.33	8.65	4.64	1.38	0.60	0.76
IL-1α	10.84	8.38	OOR <	5.33	34.15	37.99	59.27	7.76	137.57	22.72	20.83	13.32	13.01
IL-1β	26.67	8.94	3.66	4.91	25.02	37.85	16.65	4.91	134.35	13.99	1.94	11.48	7.15
IL-1RA	107.69	13.92	336.11	21.66	3.04	96.75	136.97	1.06	20.18	4.77	2.17	16.24	3.55
IL-2	0.55	OOR <	OOR <	OOR <	0.92	3.05	0.65	0.02	15.24	1.04	OOR <	0.26	0.16
IL-3	OOR <	OOR <	OOR <	0.33	0.23	0.77	OOR <	0.13	0.68	0.33	OOR <	OOR <	0.43
IL-4	0.51	0.64	2.91	1.76	1.76	6.64	1.61	1.08	4.77	0.28	OOR <	5.58	0.33
IL-5	8.44	4.56	13.40	15.41	4.94	6.07	2.54	5.86	3.33	1.27	5.59	1.65	3.88
IL-6	63.14	43.88	289.43	17.48	12.95	24.07	44.10	35.67	1.17	1.57	0.20	3.35	0.89
IL-7	21.09	3.09	3.56	7.00	12.04	8.56	7.49	0.92	5.23	0.41	0.01	1.59	2.57
IL-8	8.41	8.59	12.94	28.43	7.63	20.29	10.16	47.65	11.61	0.90	1.02	2.76	1.02
IL-9	OOR <	OOR <	13.63	5.60	8.83	19.94	2.40	OOR <	14.47	7.45	OOR <	7.57	4.28
IL-10	9.91	7.42	18.87	17.72	1.18	11.41	4.57	OOR <	6.04	0.44	OOR <	10.56	0.99
IL-12p40	41.63	21.81	126.37	57.24	14.01	323.23	48.09	11.02	191.52	68.11	56.33	38.84	102.79
IL-12p70	0.88	0.34	0.49	1.52	4.59	6.30	0.65	1.12	49.97	2.09	1.92	0.49	0.42
IL-13	95.78	8.02	61.07	11.63	45.40	66.91	72.10	23.88	50.58	34.48	29.32	54.51	OOR <
IL-15	13.92	4.11	9.96	6.45	11.57	51.98	13.92	5.81	15.74	8.04	2.94	4.11	3.04
IL-17A	9.30	17.03	1.43	6.69	12.40	13.41	0.72	1.65	33.97	4.53	0.18	6.42	2.51
IL-17E/IL-25	OOR <	OOR <	437.57	262.31	1039.24	1578.39	314.38	429.60	2018.20	476.96	OOR <	1228.47	189.18
IL-17F	2.46	1.54	3.78	31.88	8.36	110.32	29.62	2.91	95.57	21.64	1.78	8.36	OOR <
IL-18	37.66	4.82	27.73	24.67	16.56	31.84	30.88	0.23	25.15	16.52	15.71	49.68	38.79
IL-22	42.70	OOR <	51.71	OOR <	33.17	89.04	37.52	OOR <	51.71	52.74	OOR <	24.42	3.39
IL-27	4477.48	4134.36	1878.47	3564.24	254.09	3266.53	2729.04	81.55	1531.72	1540.18	1776.99	2078.08	990.15
IP-10	2330.69	154.93	492.40	791.77	319.68	330.96	25093.63	71.52	65.98	25.66	99.57	38.35	70.97
MCP-1	364.46	521.49	504.26	387.63	81.27	113.53	389.16	358.01	47.33	80.11	112.89	411.32	242.60
MCP-3	24.82	8.94	22.00	10.90	19.63	29.51	26.15	8.78	17.91	16.45	10.62	9.88	9.73
M-CSF	232.14	40.25	281.89	267.63	1494.36	599.13	443.29	227.43	108.10	74.34	33.79	16.88	129.77
MDC	415.93	688.93	483.93	254.75	351.77	237.68	325.16	315.86	294.45	437.50	326.96	110.78	394.18
MIG/CXCL9	1211.17	1812.58	7877.86	5381.92	1942.65	3703.33	3962.01	1863.35	1130.89	639.05	1130.89	792.35	1281.96
MIP-1α	23.38	16.02	8.82	20.80	32.43	61.16	23.80	20.35	56.73	28.12	8.07	18.02	13.87
MIP-1β	21.11	52.77	35.21	89.99	25.28	78.73	77.77	47.59	25.69	23.33	12.70	37.89	17.24
PDGF-AA	1344.27	1797.14	2209.35	1998.15	2165.45	422.29	1583.75	2139.64	43.81	59.64	49.49	120.45	304.98
PDGF-AB/BB	10888.77	16226.61	16510.31	14480.42	18870.15	12877.04	13524.11	22634.99	1960.23	2214.42	2088.39	3461.50	6909.87
RANTES	638.74	952.85	741.83	717.88	1502.54	1106.48	1273.18	1263.53	367.38	665.00	627.87	502.60	1086.69
TGFα	3.86	9.37	15.30	17.75	5.42	41.99	5.80	7.17	39.85	7.17	1.91	1.28	0.89
TNFα	117.31	43.33	99.74	136.45	28.30	75.24	128.47	11.80	38.15	45.45	39.98	29.54	23.30
TNFβ	11.84	6.76	10.38	9.65	12.86	8.20	18.89	4.03	8.78	5.25	2.48	1.64	1.71
VEGF-A	183.56	164.20	142.37	362.70	389.47	164.55	270.54	167.55	19.20	1.04	0.55	16.81	4.77

Note: OOR < stands for Out Of Range (below standard curve)