# Identifying host entry factors underlying respiratory virus tropism for ciliated cells

Deimante Ramonaite

A thesis submitted for the degree of Master of Science by Dissertation

Supervisors: Efstathios Giotis, Metodi Metodiev, Greg Brooke

School of Life Sciences

University of Essex

April 2026

# Abstract

Respiratory viruses pose a major global health challenge, exhibiting strong tropism for the airway epithelium, with ciliated cells often serving as primary hosts. However, the host factors driving this specificity remain incompletely understood. Human coronavirus HKU1 (HCoV-HKU1) is of particular interest due to its poorly characterised entry mechanisms and difficulty to cultivate in standard immortalised cell culture systems.

This study aimed to identify host factors enriched in respiratory ciliated cells that may influence viral entry using computational and experimental approaches. Publicly available single-cell RNA sequencing datasets from healthy adult respiratory tissues were analysed to identify transcripts enriched or depleted in ciliated epithelial cells. Candidates were selected based on expression patterns, subcellular localisation, and known virus-host interactions. CDHR3, CIB1, and EZR emerged as leading candidates and were evaluated using *in silico* docking and gain-of-function entry assays.

To investigate their roles across respiratory viruses, pseudovirus entry assays were performed using glycoproteins from SARS-CoV-1, SARS-CoV-2 wild-type and its variants (Alpha, Beta, Delta, BA.1, BA.2, BA.4, BA.5, XBB, XQ1.1), and influenza A virus. Extensive optimisation was also undertaken to establish an HCoV-HKU1 pseudovirus entry model. Overexpression of CDHR3, CIB1, or EZR did not significantly enhance viral entry *in vitro*. However, a statistically significant increase in entry was observed for the SARS-CoV-2 Alpha variant in CDHR3 overexpressing cells, suggesting a potential variant-specific interaction.

Overall, these findings highlight the complexity of respiratory virus entry and the challenges associated with modelling HCoV-HKU1 infection *in vitro*. This study provides a rational framework for identifying host factors involved in viral entry and highlights the importance of combining transcriptomic analysis with functional validation.

# Acknowledgments

Firstly, I would like to thank my family for their support throughout this journey, particularly during the more challenging moments, and for continually reminding me of how far I have come. I am especially grateful to my mum, Egle, for her constant encouragement.

I would also like to thank my partner, Connor, for believing in me on days when I did not believe in myself. I am grateful to my kitten, Julien, for his much-needed emotional support.

I would like to express sincere gratitude to my supervisor, Stathis, for continued guidance, support, and encouragement throughout my postgraduate studies, as well as for his assistance in calculating and applying the Jaccard index used in this thesis. I also thank my secondary supervisors, Metodi Metodiev and Greg Brooke, for their input and support.

I am grateful to my lab colleagues, Dongsheng, Tiffany, Tukur, and Riaz, for their support and guidance.

I would like to thank Shahid Khalid for his contribution to the molecular docking work included in this thesis. I would also like to thank Tom Peacock (Imperial College London) for providing the HKU1Del19 spike protein, and to Nigel Temperton (University of Kent) for providing the HEK293T/27 cell line.

Generative AI (ChatGPT, OpenAI) was used to assist with proofreading, grammar refinement, and document structure organising. All outputs were critically reviewed. No AI tools were used to generate data, interpret results, or produce scientific conclusions. All final content, analysis, and interpretations are my own work.

# Table of Contents

# List of figures

# List of tables

# Abbreviations

**A549** – human lung epithelial adenocarcinoma cells

**ACE2** – angiotensin-converting enzyme 2

**BSL** – biosafety level

**Caco-2** – human colorectal adenocarcinoma cells

**CALM1** - calmodulin 1

**CCDC78** - coiled-coil domain containing 78

**CDHR3** – cadherin-related family member 3

**CHO** – Chinese hamster ovary cells

**CIB1** - calcium and integrin binding 1

**COVID-19** – coronavirus disease 2019

**CoVs** – coronaviruses

**CTD** – carboxy-terminal domain

**DNAAF1** - dynein axonemal assembly factor 1

**DPP4** – dipeptidyl peptidase-4

**DYNLL1** - dynein light chain LC8-type 1

**E** – envelope

**EZR** - ezrin

**HA** – hemagglutinin

**HCoVs** – human coronaviruses

**HE** – hemagglutinin esterase

**HEK293T** – human embryonic kidney 293 cells, expressing the SV40 T-antigen

**HIV-1** – human immunodeficiency virus

**HLA-C** – human leukocyte antigen C

**HPRT** - hypoxanthine phosphoribosyltransferase

**HRVs** – human rhinoviruses

**HSP90AA1** - heat shock protein 90 alpha family class A member 1

**ICAM-1** – intracellular adhesion molecule 1

**KHSV** – Kaposi's Sarcoma-Associated Herpesvirus

**KLK13** – human kallikrein 13

**LRT** – lower respiratory tract

**M** - membrane

**MERS** – middle east respiratory syndrome

**MRC-5** – human lung fibroblast cells

**MS4A8** - membrane spanning 4-domains A8

**Mv1Lu** – American mink lung cells

**N** – nucleocapsid

**NTC** – no-template control

**NTD** – amino-terminal domain

**PDB** – protein data bank

**PEG** – polyethylene glycol

**PSENEN** - presenilin enhancer, gamma-secretase subunit

**qPCR** - quantitative polymerase chain reaction

**RBD** – receptor-binding domain

**RDE** – receptor-destroying enzyme

**RSV** – respiratory syncytial virus

**RT** – reverse transcription

**S** - spike

**SARS** – severe acute respiratory syndrome

**scRNA-seq** – single cell RNA sequencing

**SOC** - super optimal broth with catabolite repression

**TMC5** - transmembrane channel like 5

**TMPRSS2** – transmembrane serine protease 2

**TSPAN1** - tetraspanin 1

**UMAP** – uniform manifold approximation and projection

**URT** – upper respiratory tract

**VSV-G** – vesicular stomatitis virus glycoprotein G

**ZMYND10** - zinc finger MYND-type containing 10

# Chapter 1: Introduction

## 1.1 Respiratory viruses

Respiratory viruses, such as coronaviruses, influenza viruses, adenoviruses, rhinoviruses, respiratory syncytial virus, and others, infect upper and lower respiratory tracts and cause respiratory disease in humans (Boncristiani, Criado and Arruda, 2009). These infections can lead to mild or severe symptoms, and in some cases mortality, underscoring the importance of respiratory viruses even more (Welliver, 2003). Viruses capable of crossing interspecies barriers into humans pose a significant threat to global public health (Osterhaus, 2008). We have recently seen the threat that respiratory viruses pose, during the COVID-19 pandemic, where we lost millions of human lives, experienced economic repercussions, and increased poverty (Ciotti *et al.*, 2020).

### 1.1.1 Coronaviruses and influenza

Coronaviruses (CoVs) contain the largest known RNA genomes among viruses, and are classified as positive sense RNA, enveloped viruses (Fehr and Perlman, 2015). They are easily recognisable by their characteristic club-like projections found on their viral envelope which are the spike (S) proteins (**Fig. 1**) (Fehr and Perlman, 2015). CoVs can infect and cause disease in both humans and a range of animal hosts, primarily bats, cats, and camels (El-Sayed and Kamel, 2021). More recently, studies have been identified additional susceptible wildlife species, such as raccoons, deer mice, opossums, groundhogs, and rabbits (Goldberg *et al.*, 2024). Certain animal infecting CoV strains have shown the capacity to evolve genetically and cross species barriers, resulting in human infection through zoonotic transmission (Li *et al.*, 2023). Natural hosts are species in which a virus can be maintained and transmitted, whereas reservoir hosts sustain viral circulation in the population and may facilitate zoonotic spill over to humans, often without symptoms (Churchill *et al.*, 2016).

Ancestral strains of HCoV originated in bats and rodents before the zoonotic spill over to humans (Woo, 2023) (**Fig. 2**).

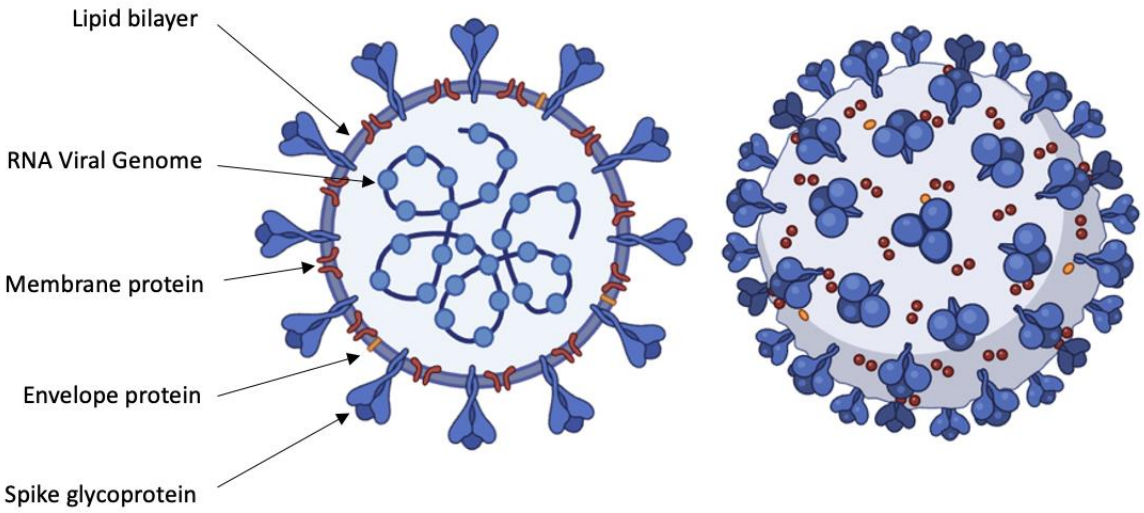



***Figure 1. Schematic structure of SARS-CoV-2.***

An illustration of the structural features of SARS-CoV-2. CoVs use these spike proteins to bind to host cell receptors, facilitating viral entry. The left image shows the viral architecture of the virus, while the right image shows a three-dimensional arrangement of the spikes on the virus surface. Image was generated using BioRender, inspired by the article 'A structural view of the SARS-CoV-2 virus and its assembly' (Hardenbrook and Zhang, 2022).

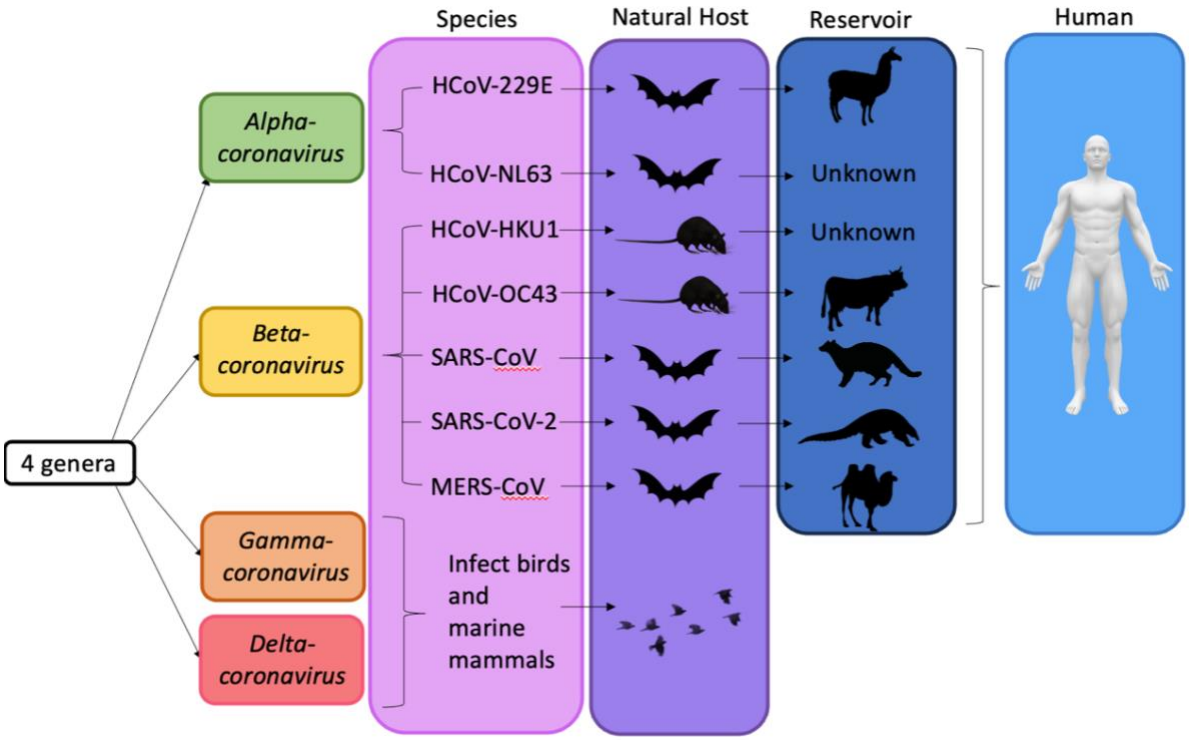


***Figure 2. The genera and zoonotic reservoirs of human coronaviruses.***

The four genera of HCoVs are depicted in the illustration: *Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus*. The *Alpha-* and *Beta-coronaviruses* include species known to infect humans. The natural hosts and reservoirs for each are illustrated, with bats identified as a common host for many human coronaviruses, alongside rodents. Animals which act as reservoirs include camelids, cattle, civets, pangolins, and camels. *Gamma-* and *Delta-coronaviruses* are not known to infect humans, they primarily infect birds and marine mammals. Figure produced using PowerPoint and information collected from the article 'Zoonotic origins of human coronaviruses' (Ye *et al.*, 2020).

The first human CoVs (HCoVs), named HCoV-229E and HCoV-OC43 were identified in the mid-1960s by Dorothy Hamre, John J. Procknow at the University of Chicago (Hamre and Procknow, 1966) and McIntosh and collaborators (McIntosh *et al.*, 1967), respectively. The late discovery of HCoVs was mainly due to the challenges in using standard cell lines to cultivate (Monto, Cowling and Peiris, 2014). However, HCoVs have been circulating for many years beforehand, as these viruses are estimated to have emerged in humans following zoonotic transmission about 268 years ago for HCoV-229E and 99 years ago for HCoV-OC43 as of a 2022 study (Forni *et al.*, 2022). In 2003 a significant discovery occurred with the identification of severe acute respiratory syndrome SARS-CoV-1, this is the first highly pathogenic HCoV and it was able to easily transmit human-to-human. This outbreak renewed the global interest in HCoVs, leading to the subsequent discovery of two more seasonalHCoVs NL63 in 2004 (Van Der Hoek *et al.*, 2004) and HKU1 in 2005 (Woo *et al.*, 2005). While NL63 and HKU1 are typically associated with causing the common cold, they have also been associated with more severe clinical conditions such as pneumonia, bronchitis, moderate-to-severe upper and lower respiratory tract infections, particularly in children, immunocompromised individuals, and the elderly (Li *et al.*, 2022). The second highly pathogenic HCoV, Middle East respiratory syndrome CoV (MERS-CoV), was identified in 2012 (Mackay and Arden, 2015). It was first isolated from a patient in Saudi

Arabia who developed severe and ultimately fatal pneumonia (Goldstein and Weiss, 2017). In late 2019, a patient in Wuhan, China, presented with atypical pneumonia, exhibiting similar clinical symptoms to SARS. Genomic analyses were run on samples from the lower respiratory tract which revealed 75% sequence homology to SARS-CoV-1. Phylogenetic studies indicated a likely bat origin with angiotensin-converting enzyme 2 (ACE2) as the entry receptor to infect human cells, similar to SARS-CoV-1 (Brant *et al.*, 2021). This novel coronavirus was called SARS-CoV-2 and the disease it causes was named the coronavirus disease 2019 (COVID-19) (*Coronaviruses*, 2024). SARS-CoV-1, MERS-CoV and SARS-CoV-2 all cause severe acute respiratory infections and are highly pathogenic, exhibiting high mortality rates. MERS-CoV and SARS-CoV-2 continue to demonstrate an increasing number of infections globally (Zhu *et al.*, 2020). In contrast, HCoV-229E, -NL63, -OC43, and -HKU1 primarily cause mild upper respiratory tract illnesses, such as the common cold, accounting for approximately 15-30% of upper respiratory tract infections (Liu, Liang and Fung, 2021).

Another respiratory virus which circulates annually and contributes to the seasonal epidemics is influenza virus. Influenza viruses are classified into four distinct types: influenza A, B, C, and D, which represent different virus species within the Orthomyxoviridae family (Krammer *et al.*, 2018). These types are differentiated based on antigenic and genetic characteristics, particularly within internal proteins such as the nucleoprotein and the matrix protein (Nuwarda, Alharbi and Kayser, 2021). Influenza virus displays seasonality with higher infection rates during colder months, showing symptoms similar to those of coronaviruses (Javanian *et al.*, 2021). While most influenza strains only cause mild symptoms and are endemic, certain variants have led to pandemics, such as the 1918 Spanish flu, which resulted in millions of deaths worldwide (Taubenberger and Morens, 2008). Influenza is a major cause of illness in humans, there are around a billion cases of infection

annually, which also includes 3-5 millions of severe cases further leading to 290,000-650,000 deaths according to WHO (WHO, 2025). Influenza A virus is an enveloped virus containing a segmented, negative-sense single-stranded RNA genome (**Fig.3**), which goes through genetic reassortment, this enhances its mutation potential (Hutchinson *et al.*, 2010). Unlike coronaviruses, which use their spike glycoprotein to bind to host cells, influenza A virus relies on its hemagglutinin (HA) glycoprotein to attach to sialic acid residues on the surface of host cells. Human strains prefer α2,6-linked sialic acid binding, whereas avian strains prefer α2,3-linked sialic acids (Leung *et al.*, 2012). This binding specificity is crucial in determining host tropism and infectivity rates. Influenza enters cells via receptor-mediated endocytosis. Once inside the endosomes, conformational change in HA is triggered due to the acidic environment, exposing the fusion peptide which then facilitates the fusion of viral and endosomal membranes (Dou *et al.*, 2018).

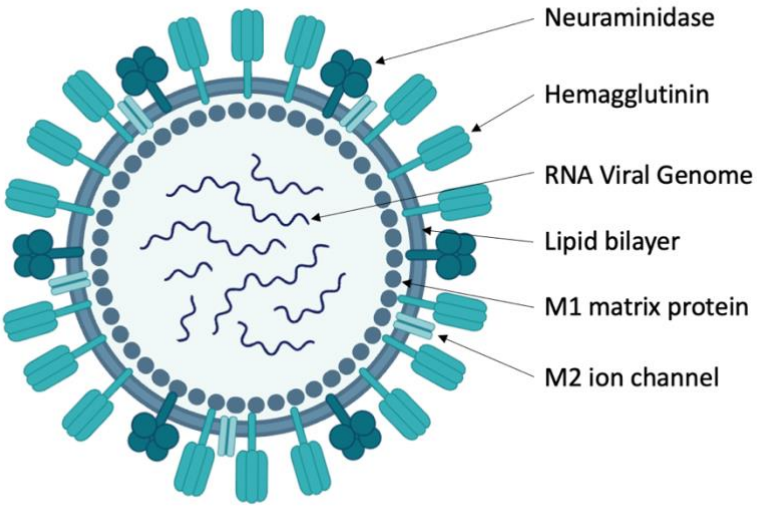



***Figure 3. Schematic structure of influenza virus.***

An illustration of the structure of influenza virus including neuraminidase, hemagglutinin, RNA viral genome, lipid bilayer, M1 matrix protein and M2 ion channel. Image was generated using BioRender, inspired by the article 'The influenza (flu) virus' (Davidson, 2015).

### 1.1.2 Adenoviruses, rhinoviruses, and respiratory syncytial virus

Adenoviruses are non-enveloped DNA viruses, human adenoviruses have over 100 serotypes, and many of them are associated with respiratory, gastrointestinal, and ocular infections (Ison and Hayden, 2016; Akello *et al.*, 2021). They infect the upper and lower respiratory tracts but usually only cause mild symptoms, and typically mainly infect children due to lack of immunity (Lynch, Fishbein and Echavarria, 2011). This virus is highly stable and spreads via a few routes: respiratory droplets, faecal-oral route, and contaminated surfaces (Jones *et al.*, 2007). Adenoviruses bind to CAR (Coxsackie and adenovirus receptor) or CD46 on the host cells, and it is internalised via clathrin-mediated endocytosis (Meier and Greber, 2004). Adenoviruses are widely used in vaccines, this is due to the ability to induce a strong immunity whilst being safe for humans as they are edited (Lasaro and Ertl, 2009). These have recently been used for vaccines against SARS-CoV-2, such as Oxford-Astra Zeneca and Johnson & Johnson (Chavda *et al.*, 2023).

There are approximately 160 human rhinoviruses (HRVs), which can cause similar symptoms to those of coronaviruses, ranging from the common cold to more serious respiratory infections. These HRVs are then further classified into three species: HRV-A, -B and -C (Lee *et al.*, 2012). Like influenza, HRVs also take the endocytic pathway, although it depends on the receptor whether they take the clathrin-dependent or -independent path (Jacobs *et al.*, 2013). The major group of HRVs, including most of HRV-A and HRV-B viruses, use the intracellular adhesion molecule 1 (ICAM-1) as their entry receptor. However, a small group from the HRV-A uses a different receptor, low-density-lipoprotein receptor (LDLR) and therefore is called the minor group. Meanwhile, HRV-C had its receptor only recognised in the last decade, which is cadherin-related family member 3 (CDHR3) (Bochkov *et al.*, 2015). CDHR3 has a single nucleotide polymorphism (rs6967330), where a guanine (G) base is changed to an adenine (A) base, and this has been associated with

increased HRV-C susceptibility in young children, which has been linked to cases of severe asthma exacerbations (Basnet *et al.*, 2019).

Respiratory syncytial virus (RSV) is the leading cause of lower respiratory tract infections in infants and young children worldwide, although it can also significantly affect the elderly and immunocompromised individuals (Shi *et al.*, 2017). RSV uses its G protein to attach to the host cells receptors, glycosaminoglycans and it uses its F protein to mediate fusion (Feldman, Audet and Beeler, 2000). Factors such as nucleolin and IGF1R have been identified to facilitating entry (Battles and McLellan, 2019). RSV reinfection is common as natural immunity is short-lived (Correa *et al.*, 2023). RSV vaccines have been licenced in 2023, specifically for the elderly and pregnant women, to protect the infant (Papi *et al.*, 2023).

### 1.1.3 Viral entry in the respiratory epithelium

When viruses enter the airway, the first barrier they encounter is the respiratory epithelium, this is a physical barrier, but it also fights against viruses in immunological functions (Pace *et al.*, 2024). To initiate entry, viruses have to interact with the cells present at the surface, this is incredibly important as viral pathogenesis depends on it (Compans and Herrler, 2007). There are two main mechanisms by which enveloped viruses enter host cells: membrane fusion and endocytosis (Yamauchi and Helenius, 2013). Despite differences in these mechanisms, the main steps of viral entry remain similar between enveloped and non-enveloped viruses (Dimitrov, 2004). The main difference is that enveloped viruses encapsulate their genome and proteins within one or more membranes which are acquired from the host when virus assembly and budding occurs (Klasse, Bron and Marsh, 1998). Viruses cannot enter every cell; cell entry is highly specific and is dictated by species-, tissue-, and cell tropism (Cossart and Helenius, 2014). Cell type is particularly important because

each expresses a distinct repertoire of receptors, which in turn determines susceptibility to viral infection. If the appropriate receptors are absent, viral tropism can be restricted (Liu *et al.*, 2021a). To initiate entry, viruses must first bind to specific host cell surface molecules known as receptors (Casasnovas, 2013). Enveloped viruses accomplish this through surface glycoproteins, such as the S glycoprotein (Cossart and Helenius, 2014). Upon receptor binding, the S glycoprotein undergoes substantial conformational rearrangements, triggered not only by receptor engagement but also by pH changes or proteolytic activation (Belouzard *et al.*, 2012). During the membrane fusion step, receptor interaction promotes the merging of viral and host membranes, enabling direct release of the viral genome into the cytoplasm. This process is often aided by host cell proteases, which cleave and activate the S protein, allowing fusion to proceed (**Fig. 4**) (Casasnovas, 2013). In contrast, during endocytosis, viruses are internalised into host cells via distinct pathways, such as clathrin-mediated endocytosis, or caveolar/raft-mediated endocytosis, mediated by the enzyme dynamin by constricting and severing the neck of the budding vesicles from the membrane (Caì *et al.*, 2015; Prichard *et al.*, 2022). Once inside, viral particles navigate through endosomal compartments before they are released from the matured endosome by fusing with the endosomal membrane, releasing their genome into the cytoplasm (**Fig. 4**) (Schelhaas, 2010). The acidic environment within the endosome often plays a crucial role in triggering fusion and genome release (Doyle *et al.*, 2024). Once inside, viruses undergo uncoating, preparing for replication and transcription (Cossart and Helenius, 2014). Human coronaviruses are the main focus of this study, hence focusing on this specific entry. Influenza is also explored in this study, and as mentioned previously it uses the HA to bind the host cell receptor, similarly to the spike of coronaviruses and it takes the endocytic pathway (Davidson, 2015).

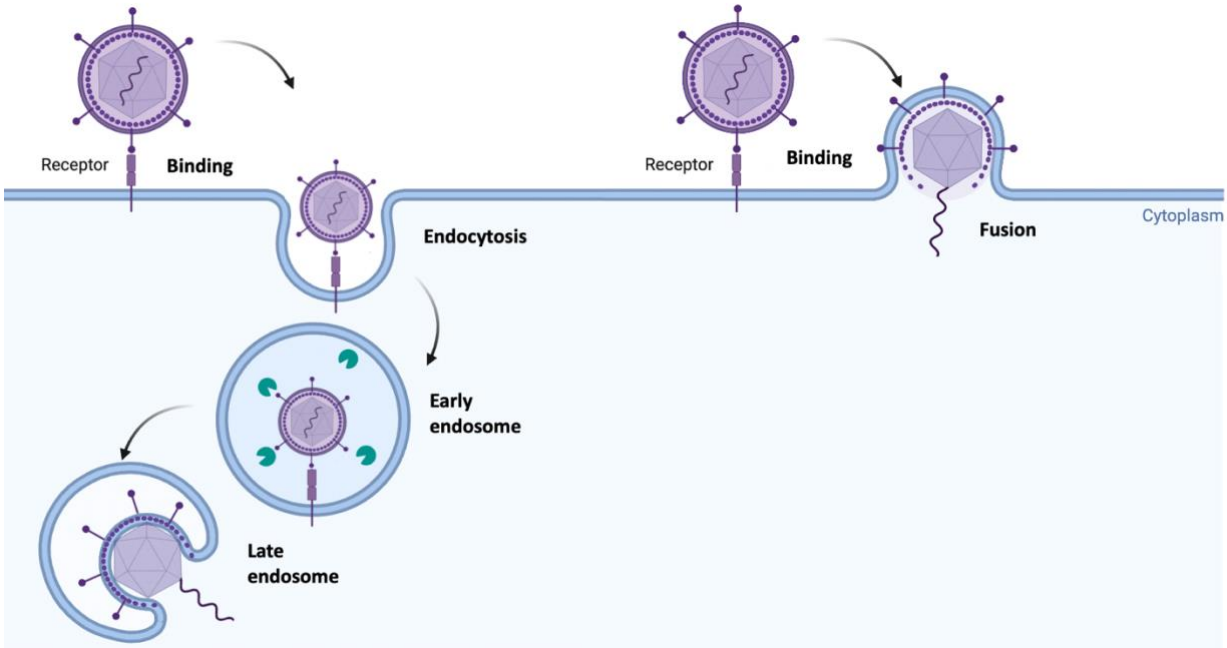



***Figure 4. Mechanism of viral entry into the host cell.***

The process begins with the virus binding to specific receptors on the host cell membrane, initiating one of two pathways: endocytosis or fusion. In the endocytosis pathway, the virus is internalised into an endosome, where maturation triggers viral uncoating. In the fusion pathway, the viral envelope directly fuses with the host cell membrane. In both cases, the viral genome is ultimately released into the cell. This figure highlights the key stages of viral entry pathways. Image was generated using BioRender, inspired by the article 'Endocytosis by Viruses and Bacteria' (Cossart and Helenius, 2014).

## 1.2 Ciliated vs non-ciliated cells in the airways

The respiratory system plays a vital role in defending against pathogens, allergens, and debris, as it clears them in order to prevent inflammation or any other illnesses, thereby maintaining homeostasis (Kia'i and Bajaj, 2019). Each cell type found in the respiratory epithelium (**Fig. 5**) has a unique function. Basal cells regenerate the epithelium, they are the main stem cells; goblet cells are the main mucus producing cells, they work together with ciliated cells and club cells on mucociliary clearance; club cells secrete uteroglobin, which has anti-inflammatory and immunomodulatory effects; neuroendocrine cells release neuropeptides that function as signalling molecules; tuft cells produce several biological

molecules, though their functions are not yet fully understood; ionocytes help regulate the ion and fluid homeostasis (Davis and Wypych, 2021; Schneider, O'Leary and Locksley, 2019; Okuda and Gentzsch, 2024). Upper respiratory tract (URT) includes the nose, nasal cavity, sinuses, pharynx, and larynx. URT is the first barrier to pathogens and its function is to filter and humidify the air, meanwhile the lower respiratory tract (LRT) contains the trachea, bronchi, bronchioles, and alveoli and is primarily responsible for gas exchange and maintaining the airway (Man, de Steenhuijsen Piters and Bogaert, 2017). The UTR is rich in goblet and ciliated cells and commonly experiences infections, especially viral ones. The LRT is not as rich in ciliated cells, and infection is harder to reach here although it can be much more severe and life-threatening (Whitsett and Alenghat, 2015; Man, de Steenhuijsen Piters and Bogaert, 2017).

### 1.2.1 Viral targeting and susceptibility to viral infections of ciliated cells

Ciliated epithelial cells serve as a primary site of infection for many respiratory viral viruses. Seasonal influenza A viruses (e.g., H1N1, H3N2) preferentially infect ciliated cells (van Riel *et al.*, 2010), while highly pathogenic avian strains such as H5N1 also infect these cells but exhibit broader tropism within the LRT (Matrosovich *et al.*, 2004). Several seasonal coronaviruses, including HCoV-HKU1, HCoV-OC43, HCoV-229E, demonstrate strong tropism for ciliated airway epithelial cells. Rhinoviruses such as HRV-C, respiratory syncytial virus, and human parainfluenza viruses also commonly infect ciliated cells (Horton *et al.*, 2025). In contrast, emerging coronaviruses such as SARS-CoV-2 infect ciliated cells but additionally target other epithelial cell types, including secretory and alveolar cells (Sungnak *et al.*, 2020; Chu *et al.*, 2020). This preference is thought to stem from the presence of specialised receptors (or cell entry factors) unique to ciliated cells, or cell host factors that facilitate virus entry or allows viral replication (Griggs *et al.*, 2017). These cells display many

specifically localised receptors on their surface, this includes receptors such as ACE2, alongside host proteases including transmembrane serine protease 2 (TMPRSS2), which facilitate SARS-CoV-2 entry by activating the spike protein (Buqaileh *et al.*, 2021; Ahn *et al.*, 2021). TMPRSS2 is an enzyme involved in facilitating viral entry by proteolytically cleaving and therefore activating envelope glycoproteins (NCBI, 2025). In a similar way, CDHR3 is exclusively expressed on ciliated cells and acts as a receptor for HRV-C, this explains the narrow cell tropism (Griggs *et al.*, 2017). Influenza viruses bind to sialic acids, which are highly expressed on the surface of ciliated cells (Luo, 2011). Human influenza A viruses preferentially bind α2,6-linked sialic acids, which are enriched in the URT, whereas avian influenza A viruses preferentially bind to α2,3-linked sialic acids, which are more abundant in the LRT (Kumlin *et al.*, 2008). Human influenza strains have efficient transmission, whereas avian strains retain receptor specificity that favours deep lung infection and increased pathogenicity but reduced transmissibility in humans (Thompson and Paulson, 2021). Upon infection of ciliated cells, viruses exploit them for replication, which in turn leads to ciliary damage and impaired mucociliary clearance, therefore evading innate immunity defence within the body (Dai, Xu and Li, 2024).

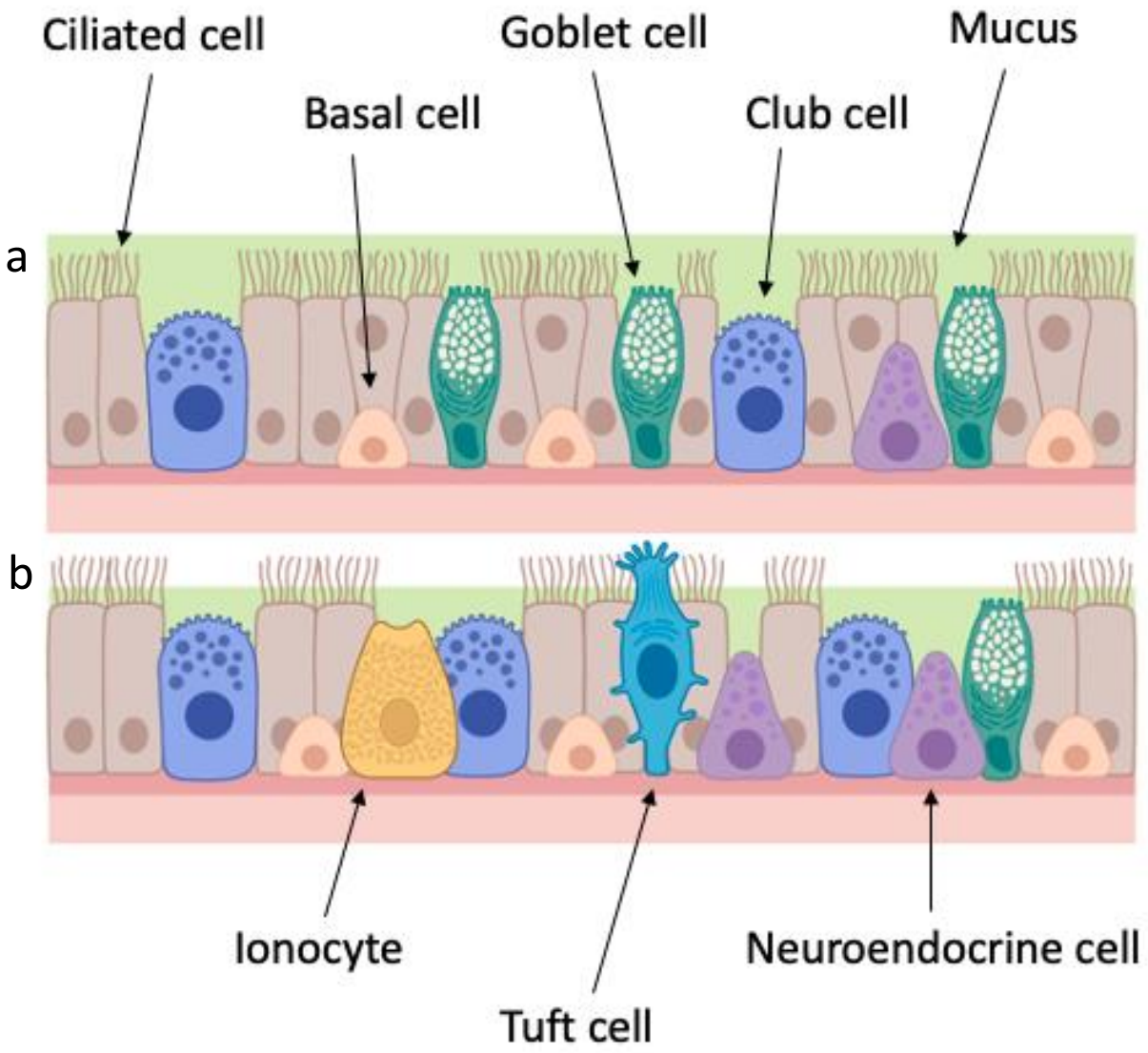



***Figure 5. The structure of airway epithelia.***

(a) Upper respiratory tract. (b) Lower respiratory tract. Basal cells regenerate the epithelium, they are the main stem cells; goblet cells are the main mucus producing cells, they work together with ciliated cells and club cells on mucociliary clearance; club cells secrete uteroglobin, which has anti-inflammatory and immunomodulary effects; neuroendocrine cells release neuropeptides that function as signalling molecules; tuft cells produce several biological molecules, though their functions are not yet fully understood; ionocytes help regulate the ion and fluid homeostasis. Image was generated using BioRender using information from the article 'Mucociliary respiratory epithelium integrity in molecular defence and susceptibility to pulmonary viral infections' (Adivitiya *et al.*, 2021).

### 1.2.2 HCoV-HKU1: entry mechanisms, the knowns, and unknowns

HKU1 contains five structural proteins: spike (S), membrane (M), hemagglutinin esterase (HE), nucleocapsid (N), and envelope (E) (**Fig. 6**). The S glycoprotein plays a crucial role in receptor recognition, viral attachment, and host cell entry, and it has remained relatively conserved across all CoVs over many years (Huang *et al.*, 2020). During entry and following binding to the receptor(s), the S protein undergoes cleavage into two subunits, S1 and S2, each with distinct functions. The S1 subunit, located at the tip of the S, is responsible for receptor binding, while the S2, found at the base of the S, facilitates membrane fusion. The S1 subunit is further subdivided into two domains: the amino (N) – terminal domain (NTD) and the carboxy (C) – terminal domain (CTD). Studies have demonstrated that HKU1 CTD serves as the receptor-binding region as it contains the receptor-binding domain (RBD). Whilst the NTD aids in the initial attachment to host factors such as 9-*O*-acetylated sialic acids (Kirchdoerfer *et al.*, 2016). A unique feature within the S1/S2 junction is the furin-cleavage site, which is not present in all coronaviruses but enhances viral entry by facilitating efficient spike cleavage (Lubinski and Whittaker, 2023). Additionally, CoVs possess an S2' cleavage site that is processed before the furin site, further contributing to improved viral infectivity (Kirchdoerfer *et al.*, 2016). The S2' subunit contains a fusion peptide along with heptad repeats 1 and 2, which are critical for membrane fusion (Neerukonda *et al.*, 2025). Recently it has been discovered that the NTD interacts with host cell glycan receptors, inducing a conformational change resulting in the exposure of the RBD. Therefore, allowing the entry receptor to bind and initiate viral entry (Yang *et al.*, 2024).

For a long time, the role of HKU1 HE protein and its interactions with 9-*O*-acetylated sialic acids remained unclear. However, a 2015 study found that the S protein of HKU1 utilized 9-*O*-acetylated sialic acids as initial cell attachment factors, also referred to as receptor determinants (Huang *et al.*, 2015). In a more recent study, it was discovered that when sialoglycans (glycans containing sialic acids) present on the host cell surface bind to the virus, they induce a conformational change at the NTD. This change facilitates the opening of the RBD enabling the host cell receptor to now be recognised (Wang *et al.*, 2024a). Additionally, the HKU1 HE protein has been observed to mediate sialate-*O*-acetylesterase receptor-destroying enzyme (RDE) activity. Although the HE protein does not directly bind to 9-*O*-acetylated sialic acids, this enzymatic activity is specific to the sialic acids recognised by the S protein (Huang *et al.*, 2015). HKU1 HE acts as an RDE by regulating the binding and release of viral progeny from the host cell. Furthermore, it reduces binding efficiency by mediating the sialate-*O*-acetylesterase RDE activity, which, in turn, decreases the exposure of the RBD to its receptor and consequently limits the binding and release of viral progeny (Wang *et al.*, 2024a).

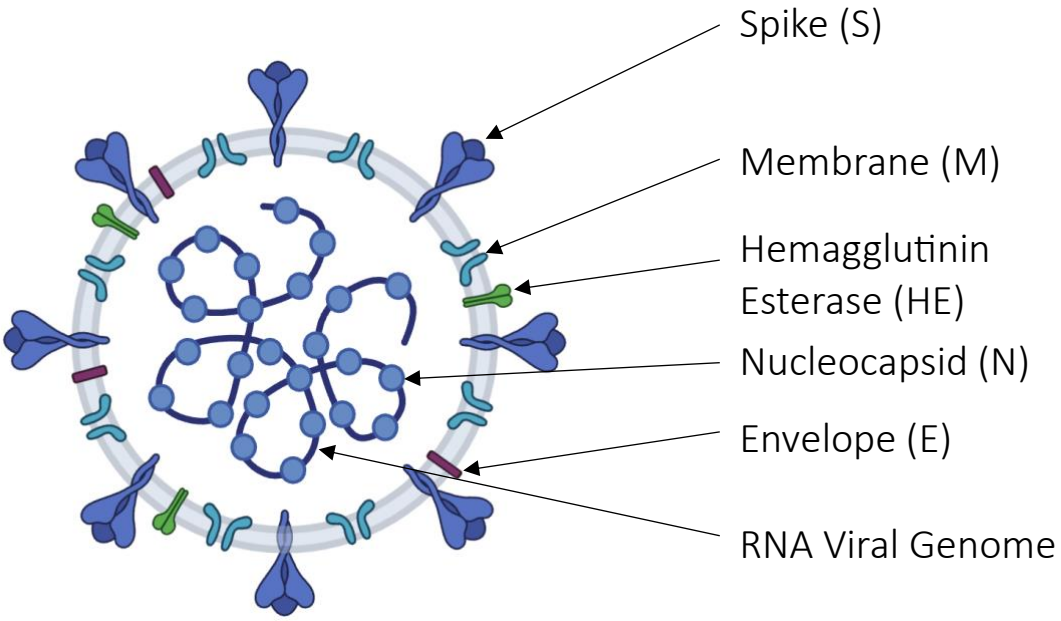



***Figure 6. Schematic representation of HCoV-HKU1.***

Showing key structural components, including the spike (S), membrane (M), hemagglutinin esterase (HE), nucleocapsid (N), envelope (E), and the RNA viral genome. Image was generated using BioRender.

Human Kallikrein 13 (KLK13) has been identified as an important factor in viral entry for HCoV-HKU1 (Milewska *et al.*, 2020). Experiments using pseudoviruses containing KLK13 demonstrated successful viral entry, although at low levels. However, when entry was inhibited using camostat and SPINK6 (both are known inhibitors of various KLKs) along with a specific KLK13 inhibitor, viral entry was significantly reduced (Milewska *et al.*, 2020). Despite its role in S glycoprotein cleavage, KLK13 alone is not sufficient for complete viral entry. Studies suggest that KLK13 serves as a priming enzyme, facilitating further cleavage at the S2' site by additional host proteases. This additional processing step is likely required to enable full viral entry and infection (Milewska *et al.*, 2020).

In late 2023, TMPRSS2 has been described as a functional entry receptor for HKU1, enabling viral entry through proteolytic activation of the spike protein (Saunders *et al.*, 2023). TMPRSS2 belongs to a family of TMPRSS, which is a subgroup of type II transmembrane serine proteases comprising 14 enzymes in humans. These membrane-anchored enzymes are synthesised as inactive zymogens and activated through proteolytic cleavage, after which they regulate a wide range of physiological processes including tissue homeostasis, proteolytic cascades, development, iron metabolism, and cancer progression as well as their involvement in viral fusion (Escalante, Wang and Ferguson, 2021). However, TMPRSS2 appears to play a dual role in HKU1 infection, acting both as a host entry factor and as a protease that cleaves and activates the S protein (Wang *et al.*, 2024a). HKU1 utilises its C terminal domain of the S protein to engage with the receptor, as the RBD is suspected to be located between amino acids 547-574. This interaction triggers a cascade leading to viral entry, either through fusion or endocytosis (Chen *et al.*, 2024). Before the dual role of TMPRSS2 was discovered, another study has suggested that human leukocyte antigen C

(HLA-C) molecule was an attachment factor for HKU1 infection and even hypothesised that it could potentially facilitate entry (Chan *et al.*, 2009). However, this hypothesis was disproven a year later when researchers tested an antiserum against HLA-C and found that it did not attenuate HKU1 infection. This demonstrated that HLA-C is not necessary for viral infection (Pyrc *et al.*, 2010).

Since the discovery of HKU1 in 2005, numerous attempts have been made to cultivate HKU1 in standard cell lines, but with little success. While other coronaviruses have been successfully propagated in standard cell lines, a 2010 study focused specifically on HKU1 found that previously utilised cell lines for virus propagation, including MRC-5 (human lung fibroblast cells), Caco2 (human colorectal adenocarcinoma cells), and A549 (human lung epithelial adenocarcinoma cells), failed to support HKU1 replication (Pyrc *et al.*, 2010). In contrast, the same study successfully infected primary airway epithelial cells, where HKU1 was observed to specifically target ciliated cells while the infection did not reach basal cells. These localisation results were very similar to the results they had with SARS-CoV-1 (Pyrc *et al.*, 2010). Since this study, there have been no further reports of successful cultivation of HKU1 in standard cell lines so far, even though the receptor for HKU1 has been discovered (Saunders *et al.*, 2023).

### 1.2.3 SARS-CoV-2 variants: known receptors and co-factors

SARS-CoV-2 is the causative agent of the COVID-19 pandemic, it primarily targets the respiratory tract and it has undergone significant genetic evolution since its emergence (Harvey *et al.*, 2021). Virus variants are strains that have one or more mutations within their genetic sequence compared to the original virus, this may result in different functional properties (Divala, 2023). Often, these mutations are advantageous to the virus, for example, enhancing transmission rates. This was specifically seen with SARS-CoV-2, where variants

of concern emerged successively: Alpha, Beta, Gamma, Delta, and Omicron (Carabelli *et al.*, 2023). These variants carried mutations in the spike and non-spike regions, ranging anywhere from a few to over 50 mutations, as seen in the Omicron variant (Farhud and Mojahed, 2022). It is believed that the Omicron subvariant evolved within an immunocompromised individual who experienced prolonged SARS-CoV-2 infection for many months, leading to adaptation of the virus to the human host further (Farhud and Mojahed, 2022). Although ACE2 is established as the principal entry receptor for all SARS-CoV-2 variants, the role of TMPRSS2 in spike activation is not universal and appears to vary with cell type, protease availability, and viral variant, highlighting the involvement of additional host entry receptors or factors (Cambi, Koopman and Figdor, 2005). For instance, C-type lectins interact with a wide range of pathogens and may promote viral entry (Cambi, Koopman and Figdor, 2005). CD147 was initially proposed as the entry receptor for SARS-CoV-1 and SARS-CoV-2 due to increased infection rates with higher CD147 levels but was eventually ruled out as a primary receptor due to the inability to bind the spike protein. Although it can still be considered an entry co-factor (Jackson *et al.*, 2022). SARS-CoV-2 primarily infects ciliated cells, with over 80% of these becoming infected within 24-48 hours post infection (Wu *et al.*, 2023). While the original SARS-CoV-2 strain already has a strong affinity for ciliated cells, emerging variants have demonstrated even greater infectivity of ciliated cells (Shrestha *et al.*, 2022). For example, Omicron variants such as BA.1, BA.5, and XBB, have shown an increased tropism to ciliated cells specifically (Parry *et al.*, 2025). Omicron mutations have altered virus-receptor binding, resulting not only in increased infectivity but also in increased transmission. Hence, the Omicron variant became the most prevalent strain globally, replacing Delta in 2022 (Wu *et al.*, 2023).

## 1.3 Studying entry at single cell level

Studying infections at single cell level can reveal several factors that would otherwise remain unknown or poorly understood. This technique reveals cellular heterogeneity during infections and highlights individual differences in response to bacterial, fungal, or viral pathogens (Huang, Wang and Yao, 2021). Such heterogeneity is usually masked when analysing bulk cell populations, giving single-cell techniques a distinct advantage in the ability to reveal differences between individual cells that might be hidden when looking at a whole population of cells (Bost and Drayman, 2024). scRNA-seq also facilitates the analysis of host-virus interactions within individual cells, revealing how different cell types respond to infection and which antiviral pathways are activated (Kelly *et al.*, 2022). In this study, analysis of the respiratory tract scRNA-seq datasets was used to identify viral entry factors expressed by ciliated cells.

### 1.3.1 Approaches for entry factor profiling: single-cell RNA sequencing

scRNA-seq is a relatively new technique, but it has already been widely used for entry factor profiling by many scientists (Jovic *et al.*, 2022). This allows us to compare transcript expression differences across individual cells, therefore providing high-resolution insights into cellular diversity (Buettner *et al.*, 2015). Furthermore, scRNA-seq reveals cellular heterogeneity within tissues and helps identify rare cell types which otherwise might be undetected (Buettner *et al.*, 2015). This technology has generated detailed atlases of the human airway epithelium, which revealed the diverse array of airway cell types (Deprez *et al.*, 2020). scRNA-seq has been used to map a wide range of tissues, such as the respiratory epithelium, this identifies which cell types are present in each of the tissues and what transcripts they express (Coate *et al.*, 2020). Allowing further investigation into which entry-related factors these cells might express. In virology, one major use of scRNA-seq has been

mapping the expression of viral entry factors, such as TMPRSS2 and ACE2 in SARS-CoV-2. This allowed researchers to assess tissue susceptibility to SARS-CoV-2 infection (Singh, Bansal and Feschotte, 2020; Qi *et al.*, 2021). Nasal goblet and ciliated cells exhibit high expression of TMPRSS2 and ACE2, they are thought to be the initial target cells of infection (Sungnak *et al.*, 2020). Further in the lower respiratory tract, alveolar type II cells have been identified to express TMPRSS2 and ACE2 as well, which means they are susceptible to SARS-CoV-2 infection and may explain the severe pulmonary symptoms in some patients (Qi *et al.*, 2021). Beyond the respiratory tract, the expression of TMPRSS2 and ACE2 is also observed in other tissues, such as the brain, gallbladder, fallopian tubes, and others, suggesting that SARS-CoV-2 may have the potential to infect multiple organs and cause diverse clinical symptoms (Qi *et al.*, 2021).

### 1.3.2 Functional validation via gain-of-function assays

Gain-of-function entry assays specifically test whether introducing or increasing specific host factors enables or enhances viral entry; these assays are widely used to identify viral receptors and other entry factors (Thevenon *et al.*, 2015). Overexpression is a powerful technique to investigate viral entry and typically involves transfecting or transducing cells with a plasmid or lentivirus vector encoding the transcript of interest to assess the encoded protein's role in viral entry (Prelich, 2012). If a normally non-permissive cell becomes susceptible to viral entry, this indicates the factor is sufficient to support entry, either as a primary receptor or co-factor (Prelich, 2012).

Pseudotyped virus systems complement this approach (see Section 1.3.3). For coronaviruses, pseudoviruses have been widely used to investigate entry mechanisms, screen entry inhibitions, and enhance susceptibility, as demonstrated in a genome-wide overexpression screen in HeLa cells for SARS-CoV-2 (Millet *et al.*, 2019; Feng *et al.*, 2023).

### 1.3.3 Pseudotyped virus systems for entry studies

Another challenge in this study, alongside the complexity of the respiratory tract, is the inability to grow some viruses in the laboratory, such as HCoV-HKU1 (Pyrc *et al.*, 2010), while others are too pathogenic and must be handled in biosafety level (BSL)-3 laboratories. In addition, viral entry needs to be studied in isolation, further limiting experimental accessibility. To overcome these challenges, researchers often utilise model viral systems such as pseudoviruses. A pseudovirus is an engineered viral particle which aims to represent the original virus, it is recombinant, and it contains the basic structure of the virus whilst its backbone is obtained from a different virus. It retains the ability to infect cells for one cycle of replication, without producing any progeny virions (Li *et al.*, 2018). This feature makes pseudoviruses a safe alternative to wild-type viruses for studying viral entry mechanisms.

Several pseudovirus systems are commonly used in virology, including lentiviral-, vesicular stomatitis virus (VSV)-, and murine leukaemia virus (MLV)-based platforms (Trischitta *et al.*, 2024). In the case of coronavirus pseudoviruses, the human immunodeficiency virus (HIV-1)-derived lentiviral particle is typically used for the backbone, which expresses packaging proteins and signals while lacking the envelope gene, as it has been deleted (Chen and Zhang, 2021). Pseudoviruses generated using the HIV-1-based lentiviral system are produced by co-transfecting packaging cells with plasmids encoding the S glycoprotein of the virus, Gag-pol (which in this case is the HIV-1), and a reporter construct such as the pCSFLW firefly luciferase plasmid, which emits luminescence upon successful infection, measured using a luminometer (Chen and Zhang, 2021). The use of pseudoviruses therefore allows viral entry to be studied in BSL-2 laboratories and provides a safe and effective approach for studying viral infection and entry mechanisms without the risks associated with handling infectious virus (Xiang *et al.*, 2022). This is particularly

advantageous for viruses like HKU1, which cannot be cultivated in standard immortalised cells lines, thereby overcoming a major limitation in coronavirus research.

Pseudotyped virus systems are also well suited for functional validation of candidate entry factors identified through transcriptomic analysis. Overexpression of candidate transcripts in permissive or non-permissive cell lines allows assessment of whether these factors influence viral entry, either as primary receptors or as co-factors (Walls *et al.*, 2020). In addition to gain-of-function approaches, receptor involvement can be further studied using inhibition or gene knock-out strategies to determine whether viral entry is reduced (Shang *et al.*, 2020a).

Extensive research on SARS-CoV-2 has highlighted a range of methodologies for identifying viral entry factors, many of which could inform studies on viruses that are difficult to propagate, like HKU1. For example, ACE2 was identified as the SARS-CoV-2 entry receptor through a combination of comparative genomics, infection assays in ACE2-expressing and control cell lines, pseudovirus entry experiments, receptor-blocking studies, and molecular modelling (Zhou *et al.*, 2020; Shang *et al.*, 2020b). These complementary approaches confirmed ACE2 as the primary receptor and suggested potential zoonotic transmission. Similar strategies are challenging for HKU1 due to its poor growth in culture, limiting direct infection assays.

## 1.4 Aims and objectives for the study

Despite extensive research on respiratory viruses, significant gaps remain in our understanding of the host factors that determine viral tropism within the airway epithelium, particularly in ciliated cells. While receptors such as ACE2 and TMPRSS2 have been identified for SARS-CoV-2, and TMPRSS2 has recently been proposed as a functional receptor for HCoV-HKU1, supporting experimental evidence remains limited. In particular, it

remains unclear which ciliated-cell-enriched host factors may act as receptors or entry co-factors for viruses such as HCoV-HKU1, SARS-CoV-2 and its variants, and influenza A virus. The overarching aim of this study is to identify and functionally evaluate host factors enriched in respiratory ciliated cells that may influence viral entry. Specifically, this study aims to identify candidate ciliated-cell-enriched host factors using publicly available single-cell transcriptomic data from healthy human adult respiratory tract samples to identify host factors enriched in ciliated cells, identify candidate genes, which are found on membranous localisations and have relevance to viral entry. Finally, functionally assess with pseudotyped viruses to evaluate their impact on entry of HCoV-HKU1, SARS-CoV-1, SARS-CoV-2 wild-type and its variants, and influenza A virus, providing insight into the tropism of ciliated cells.

# Chapter 2: Materials and methods

## 2.1 Bioinformatics

### 2.1.1 Data collection

The datasets used for this study were obtained from the GEO, a publicly available RNA-sequencing database. The data used was respiratory scRNA-seq of 26 donors who were collected from 12 different studies, they were all healthy uninfected/untreated adults, which acted as controls. Samples were primarily swabbed from the nasal canal, trachea, or bronchi. There was a mixture of males and females, ages ranging from 18 to 57 years old from the known samples, some samples were anonymised so no information could be obtained (**Table 1**). The study aimed to cover both genders and a wide range of ages. No studies recorded the patient's ethnicity. Diseased samples were excluded. Further underlying conditions of the donors were not included in the studies. We focused on scRNA-seq data to look at individual cell lines present in each sample; this is not possible with bulk RNA seq data therefore it was excluded. These datasets were selected to enable comparison of transcript expression across airway cell types, with particular focus on ciliated epithelial cells. One limitation was the availability of datasets and the use of different scRNA-seq technologies. To minimise technical bias and platform-specific artefacts, we deliberately prioritised datasets generated using comparable scRNA-seq technologies, thereby reducing batch effects, and improving the robustness of cross-study comparisons.

***Table 1. Demographic characteristics of donors (age and sex) from publicly available datasets by study author.***

| Study | No. of samples | Sample | Sex | Age |
|---|---|---|---|---|
| Chan, *et al.* | 2 | Nasopharyngeal | N/A | |
| Pezzulo, *et al.* (a) | 1 | Airway epithelial | Male | 57 |
| Cheemarla, *et al.* | 1 | Bronchial | Female | 38 |
| Pezzulo, *et al.* (b) | 1 | Airway epithelial | Anonymised | |
| Paranjapye, *et al.* | 7 | Bronchial and nasal | 3 male and 4 female | 18-41 |
| Cooney, *et al.* | 4 | Airway epithelial | 3 male and 1 female | 31-49 |
| Morse, *et al.* | 4 | Mix of upper and lower lobe lung | 2 male and 2 female | 23-57 |
| Okuda, *et al.* | 2 | Upper and lower airway epithelium | N/A | |
| Sohail, et al. | 1 | Lung | Male | 46 |
| Wang, *et al.* | 1 | Airway epithelium distal | N/A | |
| Ravindra, *et al.* | 1 | Bronchial | Anonymised | |
| Dirvin, *et al.* | 1 | Tracheal | Anonymised | |

### 2.1.2 Single-cell RNA-seq data processing, annotation, and enrichment analysis

The downloaded scRNA-seq datasets were in csv., tsv., or h5 formats from 10X Genomics technology, which were analysed using Trailmaker™ (Parse Biosciences, https://www.parsebiosciences.com/data-analysis/ ). Trailmaker™ is an online platform for processing, visualisation, and analysis of single-cell transcriptomic data (Watson, 2025). Each dataset was analysed independently to account for differences in donor number and

sampling site. Data were processed using default quality-control parameters, with no manual adjustment.

Cell populations were annotated using established marker gene sets and reference-based models specific to human lung tissue, facilitating robust identification of ciliated and non-ciliated epithelial cell types (Watson, 2025). Transcript expression was analysed for individual annotated cell types, allowing direct comparison between ciliated and non-ciliated populations using transcript lists, UMAP visualisations, and violin plots. UMAP is a method used to turn complex, high-dimensional data (for example, transcript expression in cells) into a simple 2D or 3D plot. It does this by placing similar data points close together and dissimilar ones farther apart making patterns and clusters easier to see (Tjoonk, 2023). Violin plots presented the raw expression of each transcript within the different cell types. This score shows the exact count of transcripts for a specific gene detected in a cell before any normalization or log-transcription are applied.

For downstream analysis, transcripts enriched in ciliated cells were filtered based on a $\log_2$ fold-change threshold of $\geq 0.8$ for upregulated transcripts, or $\leq -1$ for downregulated transcripts. This threshold was selected to prioritise highly upregulated/downregulated transcripts. Across all datasets, this approach reduced the number of transcripts, which were exported for further localisation and prioritisation analyses.

### 2.1.3 Subcellular localisation and membrane-associated transcript prioritisation

Transcripts identified as upregulated or downregulated in ciliated cells (compared to non-ciliated cells) by Trailmaker™ were subsequently filtered based on previously known subcellular localisation using SubcellulaRVis. SubcellulaRVis is an online user-friendly tool that describes the subcellular localisation of the transcript list provided, but only of those transcripts that have previously ascertained locations (Watson *et al.*, 2022). Due to this, it

does not predict localisation for new transcripts, which have not been mapped within SubcellulaRVis, therefore it adds a limitation of not mapping all of the transcripts presented (Watson *et al.*, 2022). Each samples' filtered transcripts were run, selecting such settings: organism - human, annotation source - transcript ontology, identifier type - HGNC symbol, visualisation of whole cell. Once the enrichment was calculated and presented, the majority of the transcripts were localised. A table was outputted for a total of 170 transcripts with expression found at the plasma membrane and cytosolic-facing membranes (see Appendix, Table 1). Additionally, a list of transcripts was provided which included all the transcripts that could not be localised (see Appendix, Table 2). Transcripts encoding proteins specifically localised to the outer plasma membrane could be investigated, alongside additional membrane-associated factors that may affect viral entry. Transcripts were collected for downstream analysis if they were both highly expressed in ciliated cells (FC $\log_2 \geq 0.8$) and localised to membrane-associated compartments, resulting in a refined shortlist of candidate transcripts.

### 2.1.4 Heatmap visualisation of candidate transcript expression

Heatmaps were produced using a web server which is free to use for all users and has a wide range of applications, including expression-based data, called Heatmapper (Babicki *et al.*, 2016). Within the 'expression' tab on the website, an excel file was uploaded with all of the transcript values at $\log_2$ FC that were found to be located on the surface and cytosolic membranes within the cell, as mentioned in the section above. The scale and colouring were adjusted, scale type and clustering method selected to 'none' and the other settings were left at the automatic setting. The values start below 1 ($\log_2$ values) with a deep pink moving through lighter shades into green, with deep green having the highest expression rates. White

represents no expression. Heatmaps were used to visualise relative expression patterns of candidate transcripts across donors.

### 2.1.5 Jaccard index

Jaccard index analyses were performed in a Jupyter Notebook using Python 3.9 on Linux. Expression calls for 20 cilia-associated transcripts across 26 donors were imported into Pandas 1.3.5 and converted to a Boolean presence/absence matrix. Pairwise Jaccard similarities (NumPy 1.21.4) were calculated and transformed into distances (1 – Jaccard). We applied SciPy 1.7.3's average-linkage hierarchical clustering with optimal leaf ordering to arrange transcripts by co-expression overlap. The dendrogram and heatmap were plotted side by side in Matplotlib 3.4.3, with candidate transcripts axis labels coloured red for emphasis. All code was version-controlled with Git 2.34.

### 2.1.6 Molecular docking

HKU1 and candidate transcript structures were obtained from the Protein Data Bank (PDB). Molecular docking was performed using PyMOL software (v.3.1.0), where the HKU1 structure was uploaded, assessed, and cleaned up by removing water, sugars. The CTD domain of HKU1 was specifically identified using UniProt, and then removed and isolated in PyMOL. All 3 candidates were cleaned and prepared the same way by removing water and sugars. All of these steps were performed before HDOCK, which means the specific binding domain of interest is already specified. Each of the candidates were run in HDOCK with HKU1 CTD domain, which then outputted a list of predicted interaction points and residue pairing. Interaction points were visualised in PyMOL.

PDB provides access to 3D structures of proteins, nucleic acids, and their complexes ('Protein Data Bank: the single global archive for 3D macromolecular structure data,' 2019).

The HKU1 spike structure was analysed using PyMOL, a molecular visualisation tool that enables the examination of 3D molecular conformations, colouring, and interaction analysis. Molecular docking was therefore used as an exploratory approach to assess potential physical interactions between HKU1 spike domains and candidate host proteins.

## 2.2 Cell lines

Cell lines were cultured and subcultured using aseptic technique in a laminar flow hood in a BSL-2 laboratory following the guidelines and health and safety according to University of Essex. Cell lines used include human embryonic kidney 293 cells, expressing the SV40 T-antigen (HEK293T and HEK293T/27) (HEK293T/27 kindly provided by N. Temperton, University of Kent), human adenocarcinoma alveolar basal epithelial cells, stably expressing ACE2 and TMPRSS2 (A549+ACE2+TMPRSS2), Chinese hamster ovary cells (CHO), human colorectal adenocarcinoma cells (Caco-2), American mink lung cells (Mv1Lu). Cells were thawed/frozen as in the methods below, long term storage in liquid nitrogen. 37°C 5% $CO_2$ incubator was used to grow all cell lines. Media used for each cell line is shown in **Table 2**.

***Table 2. Complete growth medium for each cell line used.***

Cells were grown at a 37°C 5% $CO_2$ incubator and subcultured 2-3 times a week. All medium was bought from Gibco Thermo Fisher Scientific.

| **Cell lines** | **Growth media** |
|---|---|
| A549+ACE2+TMPRSS2<br>(Human adenocarcinoma alveolar basal epithelial cells, stably expressing ACE2 and TMPRSS2) | F-12K Nut Mix<br>10% FBS<br>2mM L-Glutamine (L-Glu)<br>2mg/ml Geneticin (G418) |

| | |
|---|---|
| | 200 µg/ml Hygromycin B<br>100 Units Penicillin and 100µg Streptomycin/ml |
| Caco-2 (Human colorectal adenocarcinoma cells) | Eagle's Minimum Essential Medium (EMEM)<br>10% Foetal Bovine Serum (FBS)<br>1% Non-Essential Amino Acids (NEAA)<br>2mM L-Glutamine (L-Glu)<br>100 Units Penicillin and 100µg Streptomycin/ml |
| CHO (Chinese hamster ovary cells) | F-12K Nut Mix<br>10% FBS<br>2mM L-Glutamine (L-Glu)<br>100 Units Penicillin and 100µg Streptomycin/ml |
| HEK293T and HEK293T/27 (Human embryonic kidney 293 cells, expressing the SV40 T-antigen) | Dulbecco's Modified Eagle's Medium (DMEM)<br>10% Foetal Bovine Serum (FBS)<br>2mM L-Glutamine (L-Glu)<br>100 Units Penicillin and 100µg Streptomycin/ml |
| Mv1Lu (American mink lung cells) | Eagle's Minimum Essential Medium (EMEM)<br>10% Foetal Bovine Serum (FBS)<br>2mM L-Glutamine (L-Glu) |

| | |
|---|---|
| | 100 Units Penicillin and 100µg Streptomycin/ml |

### 2.2.1 Thawing, freezing, and subculturing of adherent cell lines

For routine maintenance of adherent cell lines, cells were thawed, subcultured, and cryopreserved using standard aseptic techniques. Cryovials stored in liquid nitrogen were transported on dry ice and rapidly thawed in a 37 °C water bath until a small ice crystal remained, after which the cell suspension was diluted dropwise with complete growth medium and centrifuged to remove residual cryoprotectant before resuspension and incubation at 37 °C in a humidified 5% $CO_2$ atmosphere. Cell cultures were routinely examined by inverted microscopy to assess morphology, confluence, and contamination, and cells at approximately 80–90% confluence were subcultured by PBS washing, trypsin-mediated detachment, and reseeding into fresh complete medium. For long-term storage, adherent cells were detached, pelleted by centrifugation, and resuspended in complete growth medium supplemented with DMSO at cell line–appropriate concentrations, transferred to cryovials, gently mixed to ensure homogeneity, and frozen overnight at −80 °C in an isopropanol-based controlled-rate freezing container before transfer to liquid nitrogen for long-term storage.

## 2.3 Plasmids

The plasmids used included p8.91, which is the HIV-1 derived lentiviral vector used for the backbone of pseudoviruses (catalogue number: VC101743), the pCSFLW firefly luciferase expression vector also used for pseudoviruses to detect luciferase (catalogue number: TR30004), envelope protein VSV-G (Vesicular stomatitis virus glycoprotein G)

which was used as a positive control in pseudovirus infection (catalogue number: RG216583), pcDNA3.1 was used as a negative control as it has no gene inserted (catalogue number: V79020), TMPRSS2 was used in pseudovirus production as well as expression on the cell surface to test entry of pseudoviruses (catalogue number: SC327067), CIB1 (catalogue number: RG201591), EZR (catalogue number: SC117991), CDHR3 (catalogue number: RC218636) were used for overexpression on target cell lines as candidate transcripts that may affect entry to ciliated cells, HKU1 spike (catalogue number: 168911), HKU1Del19 spike (kind donations from Dr Peacock at Imperial College London), and SARS-CoV-2 spike (catalogue number: VC102557), and SARS-CoV-2 spike variants Wuhan, Alpha, Beta, Delta, BA.1, BA.2, BA.4, BA.5, XBB, XQ.1 were used in pseudoviruses.

### 2.3.1 Bacterial transformations and plasmid propagation

2-5µl of plasmid was added to 30-50µl *E.coli* competent cells (kind donations from Lynwen James at the University of Essex) (neg. control is *E.coli* competent cells only). Cells were incubated on ice for 30 min, heat shocked at 42°C for 45 secs and incubated on ice again for 5 min. 500µl of super optimal broth with catabolite repression (SOC) broth was added to each tube and they were placed on a shaker for 1h at 37°C at 250rpm. 50µl of each culture was added to the agar plates and incubated overnight at 37°C. The next day, the plates were inspected for colonies and either, left for longer if not enough growth is seen, sealed, and refrigerated, or a single colony was picked using a sterile loop and inoculated in 3-50ml broth/ampicillin (100µg/ml). Tubes were placed on a shaker at 37°C at 250rpm overnight. The following day, glycerol stock was made by mixing 0.5ml cells with 0.5ml sterile 50% glycerol solution, two tubes were made and stored in -80°C. Further, DNA extractions were performed using plasmid miniprep or midiprep kits.

### 2.3.2 Bacteria recovery from glycerol stocks

Tubes were retrieved from the -80°C freezer and kept on ice to ensure they did not thaw. A sterile loop or pipette tip was used to scrape some frozen bacteria from the top of the tube. This was then streaked onto an agar plate and grown in a 37°C incubator overnight. The following day the plate was checked for colonies and normal transformation protocol was followed.

## 2.4 Pseudovirus production

Pseudoviruses were generated using an HIV-1 based lentiviral packaging system expressing viral spike glycoproteins of interest (Temperton, Wright and Scott, 2015). Cells were seeded at 7-9x$10^5$ in a T25 flask, or 1.7-2x$10^6$ in a T75 flask in 4ml or 8ml of media, respectively. Cells were incubated overnight at 37°C 5% $CO_2$. Next day the DNA mix was prepared (**Table 3**), FuGENE HD was used as the transfection reagent. Once the DNA-transfection mix was ready, it was added to the cells. Cells were incubated for 48h at 37°C 5% $CO_2$. For HKU1/HKU1Del19 PVs Neuraminidase was added at the 24h point. After 48h the supernatant was harvested, filtered, and stored at -80°C, cell media was replaced. At 72h supernatant was collected and filtered again, mixed with the previous day supernatant and aliquoted. PVs were stored at -80°C.

***Table 3. DNA-Transfection mix for pseudovirus production in T25 and T75 flasks.***

| | **Spike plasmid (ng)** | **TMPRSS2 (ng)** | **p8.91 (ng)** | **pCSFLW (ng)** | **OptiMEM (µl)** | **FuGENE HD (µl)** |
|---|---|---|---|---|---|---|
| **T25** | 700 | 1100 | 700 | 1100 | 120 | 10.8 |
| **T75** | 1000 | 1500 | 1000 | 1500 | 200 | 15 |

### 2.4.1 Pseudovirus transduction assay

50µl of PV was added to an appropriate number of wells in a 96-well plate, including the same amount of the positive control (VSV-G) and negative control (containing no S plasmid). After the specific cells were counted, 40,000 cells in a total volume of 50µl were added to each well containing the PVs. These were mixed by pipetting up and down several times ensuring homogeneity. The plate was incubated for 48h at 37°C 5% $CO_2$.

### 2.4.2 Luciferase assay

Bright-Glo was thawed on ice, covered (due to being light sensitive). A 1:1 ratio of PBS to Bright-Glo was prepared. The media was removed from each well. 50µl of PBS+Bright-Glo was added to each well and incubated for 5 mins, covered. The lysate from each well was transferred to a white, opaque 96-well plate for luminescence measurement. Luminescence was measured using Promega's GloMax®-Multi Detection System.

### 2.4.3 Reverse transcription – quantitative polymerase chain reaction (RT-qPCR)

Cells expressing the candidate transcripts of interest were sub-cultured in a 96-well plate. In parallel, pseudoviruses were transduced into the same cell lines in separate wells, ensuring that all conditions remained identical up to the final step. After a 48-hour incubation period, cells transduced with pseudoviruses were assessed for luminescence. Whilst the non-transduced cells were scraped using a pipette tip, suspended in PBS, and subsequently frozen at −80 °C for later use.

RNA extraction was carried out using the RNeasy Mini Kit (Qiagen, catalogue number: 74106) according to the manufacturer’s protocol. RNA concentration and purity were determined using a NanoDrop spectrophotometer. Genomic DNA removal and cDNA synthesis were performed with the QuantiTect Reverse Transcription Kit (Qiagen, catalogue

number: 205310), following the manufacturer's instructions. Quantitative PCR (qPCR) reactions were prepared in a 10µl total volume per sample, consisting of: 5µl qPCRBIO SyGreen Blue Mix, 0.4µl forward primer, 0.4µl reverse primer (**Table 4**), 1µl cDNA, and 3.2µl RNase-free water. Each sample was run in technical duplicate. Hypoxanthine phosphoribosyltransferase (HPRT) gene was used as the housekeeping control, also in technical duplicate. No-template controls (NTCs) were included for each run, along with reverse transcription (RT) controls.

***Table 4. Primers used for qPCR.***

Primers identified using the Primer3Plus online tool, using gene sequences obtained from GEO database.

| **Gene** | **Forward primer (5'-3')** | **Reverse primer (3'-5')** |
|---|---|---|
| **CDHR3** | TGTGAAGGATGAGGTTGGTG | TCCAGGGTTTGCTCTTTCTAC |
| **CIB1** | TGCCTTCCGCATCTTTGACT | ATCTCAGACGCACTAAGCCG |
| **EZR** | TGGAGTTGATGCCCTTGGAC | GTCAGGTGCCTTCTTGTCGA |
| **TMPRSS2** | TCGATTCTTGCCAGGGTGAC | TTCCCGTACACTCCTGGTCT |
| **HPRT** | CCTGGCGTCGTGATTAGTGA | CGAGCAAGACGTTCAGTCCT |

Relative gene expression was calculated using the ΔΔCt method. Ct values of target genes (CDHR3, CIB1, and EZR) were first normalised to the reference gene (HPRT) to obtain ΔC values. ΔΔCt values were then calculated relative to the control condition. Fold change in gene expression was expressed as $2^{-}$ ΔΔCt.

## 2.5 Statistical analyses

Statistical analyses and graphs were performed and generated using GraphPad Prism (v.10.6.1). Each experiment was performed with at least 3 independent replicates. First, the data were assessed for normality using the Shapiro-Wilk normality test. As all samples passed the normality test, they were further analysed for one-way or two-way ANOVA to evaluate the statistical significance using Dunnett's multiple comparison test and Tukey's multiple comparison test, respectively. Values of the same pseudoviruses were compared between the different cell lines (p-values set at $p \leq 0.05$(*), $< 0.01$ (**), $< 0.001$ (***), $<0.0001$ (****)).

# Chapter 3: Identification of ciliated cell host factors with potential roles in respiratory viral entry

This chapter describes the identification of candidate transcripts enriched in respiratory ciliated cells from 26 healthy donors from 12 publicly available transcriptomic datasets. Analyses focused on transcript expression patterns between ciliated versus non-ciliated cells. A list of transcripts was generated, which was filtered and further analyses were performed on predicted localisation, Jaccard index, and molecular docking. These data were combined with literature review to prioritise candidates with potential roles in respiratory viral entry. Only transcripts consistently enriched across multiple donors were prioritised.

## 3.1 Identification of transcripts enriched in respiratory ciliated cells

Initially, 170 transcripts with expression levels above a $\log_2$ FC threshold of $\geq 0.8$ were selected to exclude low-abundance transcripts while retaining sufficient candidates for downstream analyses. From this set, transcripts that were most consistently enriched in ciliated cells across donor samples were prioritised, resulting in a final set of thirteen candidate transcripts. **Figure 7** summarises the thirteen transcripts most commonly enriched in ciliated cells across all datasets, with absence of expression indicated in white. The chosen transcripts were consistently overexpressed in ciliated cells across most donor samples, with only occasional absences.

Several transcripts, including CALM1, CIB1, DYNLL1, EZR, HSP90AA1, PSENEN, TMC5, TSPAN1, showed moderate enrichment in ciliated cells but were also expressed in other airway cell types (**Fig. 8 and 9**). For example, HSP90AA1 additionally showed increased expression in basal and goblet cells. However, CCDC78, CDHR3, DNAAF1, MS4A8 and ZMYND10 displayed strong and relatively specific enrichment in ciliated cells

compared to other cell types, suggesting a closer association with ciliated cells (**Fig. 8 and 9**). Transcript expression was assessed using Trailmaker™, with violin plots used to evaluate the raw expression across cell types. Raw expression represents the exact count of transcripts for a specific gene detected in a cell before any normalization or log-transcription are applied.

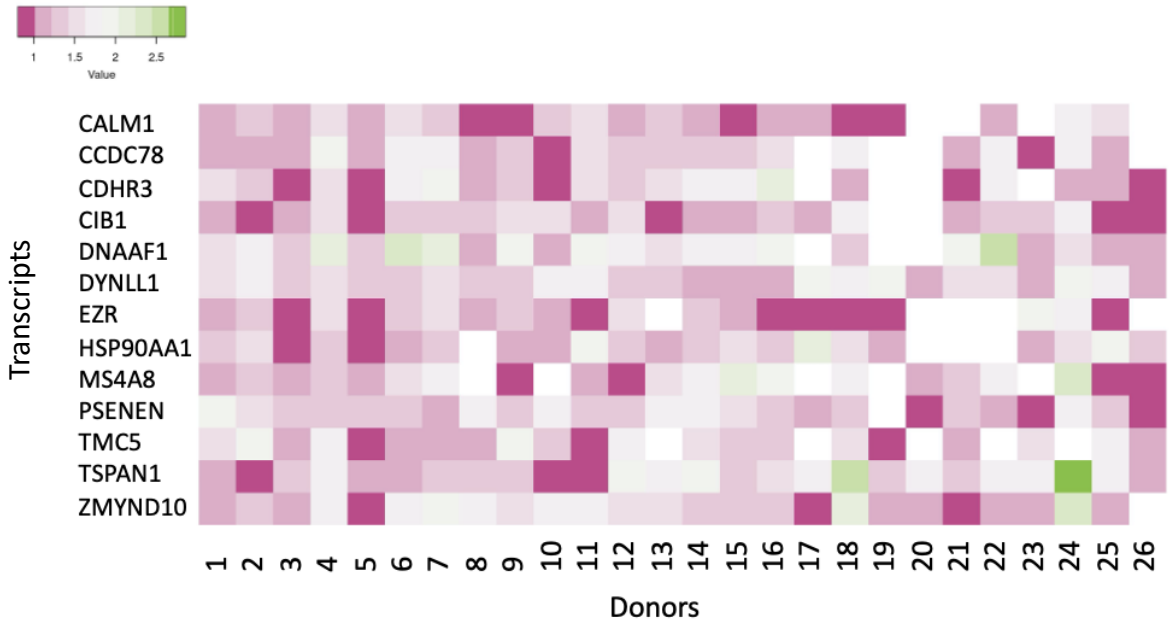
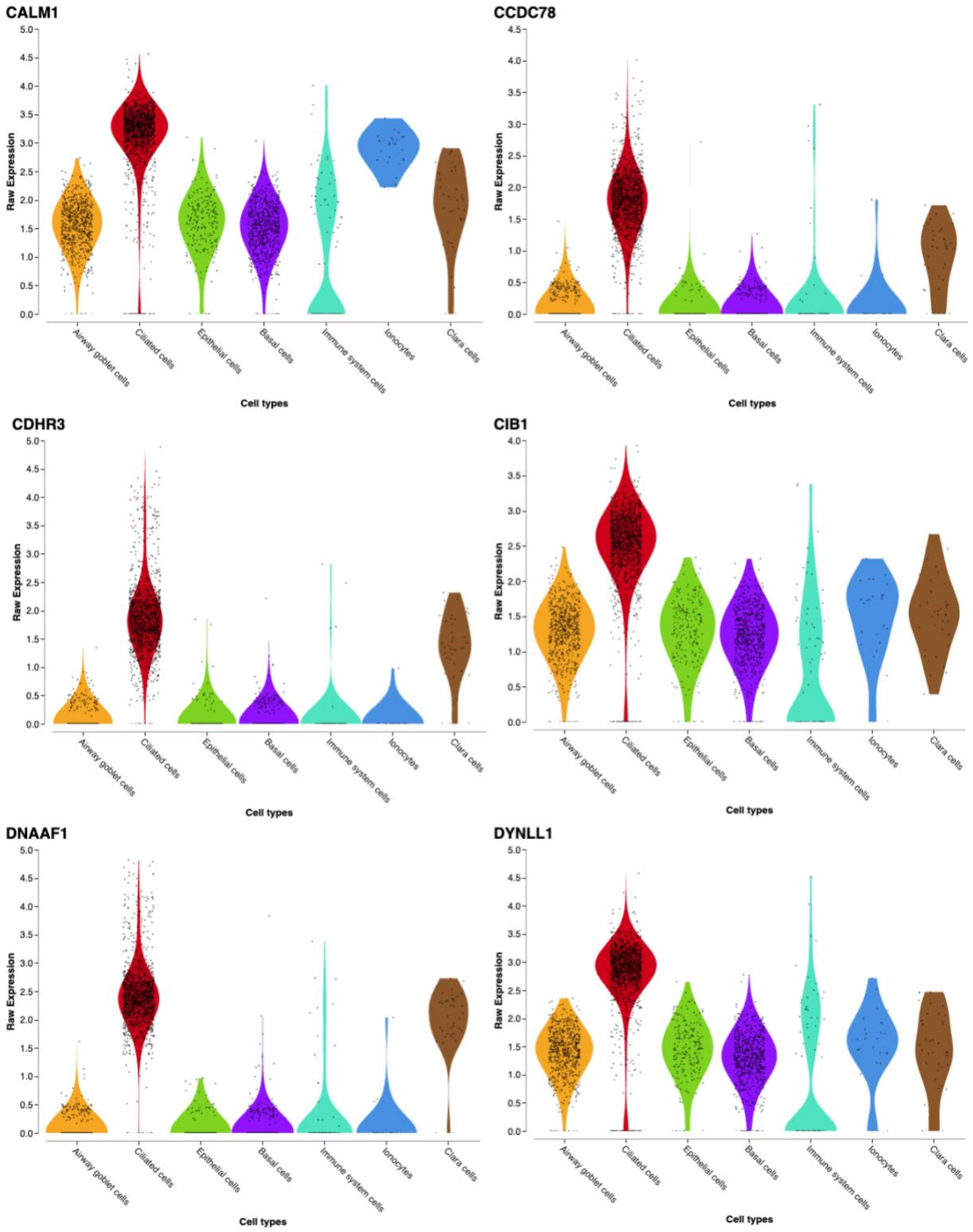



***Figure 7. Transcripts enriched in respiratory ciliated cells across donors from publicly available datasets.***

The heatmap shows the thirteen transcripts most consistently enriched in respiratory ciliated cells across twelve publicly available transcriptomic datasets from healthy airway donors. Transcript expression was analysed using Trailmaker™. Only transcripts with a $\log_2$ FC value of $\geq 0.8$ were included, representing differential expression between ciliated and non-ciliated cells in each sample. Colour intensity represents $\log_2$ FC in expression, with higher expression indicated in green and lower expression indicated in pink. White indicates absence of detectable expression in the corresponding donor sample.


CALM1
CCDC78
CDHR3
CIB1
DNAAF1
DYNLL1
Raw Expression
Cell types
Airway goblet cells
Ciliated cells
Epithelial cells
Basal cells
Immune system cells
Ionocytes
Clara cells

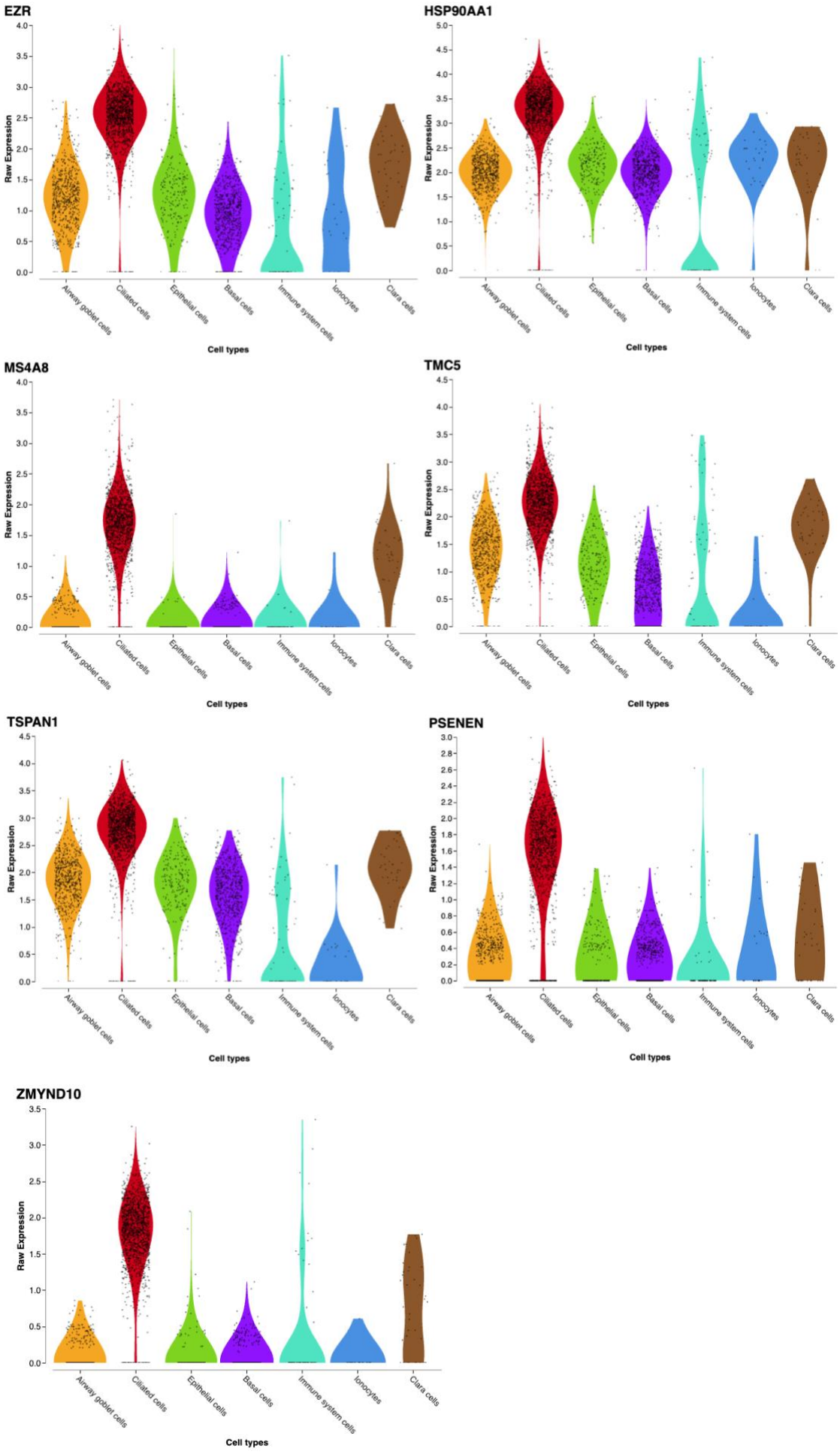

EZR
HSP90AA1
MS4A8
TMC5
TSPAN1
PSENEN
ZMYND10
Raw Expression
Cell types
Airway goblet cells
Ciliated cells
Epithelial cells
Basal cells
Immune system cells
Ionocytes
Clara cells

***Figure 8. Transcript expression of CALM1, CCDC78, CDHR3, CIB1, DNAAF1, DYNLL1, EZR, HSP90AA1, MS4A8, TMC5, TSPAN1, PSENEN and ZMYND10 across airway cell types.***

Violin plots showing raw expression of the three transcripts across airway cell types from donor 6 generated using Trailmaker™ (Paranjapye *et al.*, 2022). Each dot represents an individual cell. Raw expression represents the exact count of transcripts for a specific gene detected in a cell before any normalization or log-transcription are applied. All transcripts show highest and most concentrated expression in ciliated cells compared to other airway cell populations.

Additional analyses were performed to identify transcripts that were consistently downregulated in ciliated cells, using $\log_2$ FC threshold of ≤ -1. **Figure 11** shows the most commonly downregulated transcripts with absences seen in white. While the present study focused on transcripts upregulated in ciliated cells, downregulated transcripts remain of potential interest for future investigation.

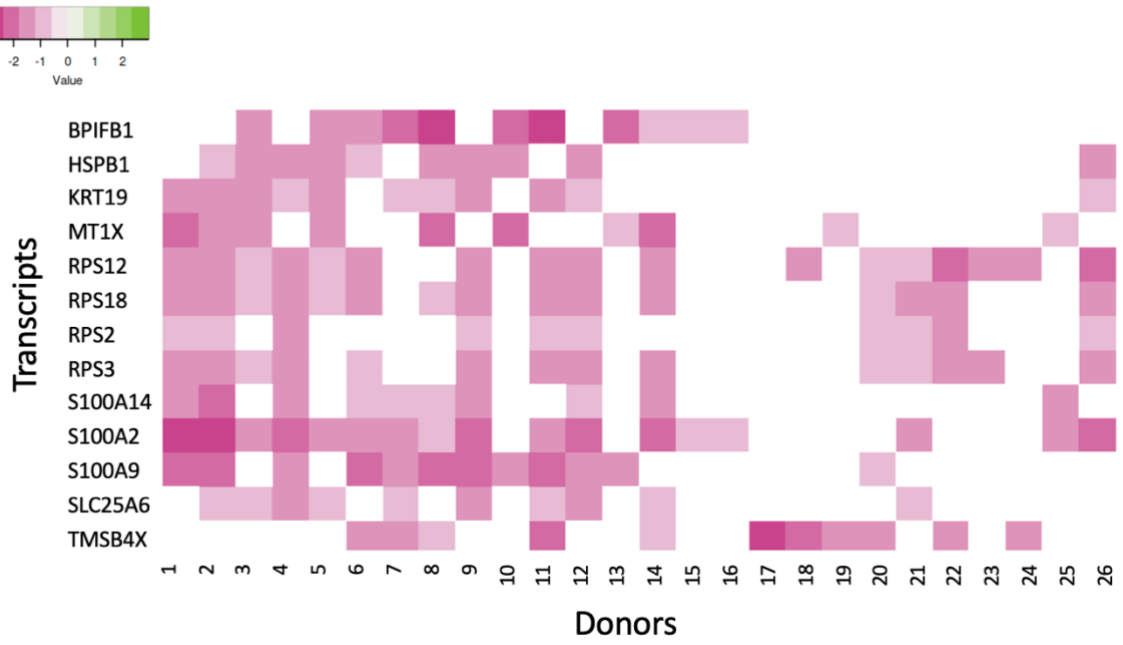



***Figure 9. Transcripts downregulated in respiratory ciliated cells across publicly available airway datasets.***

The heatmap shows the thirteen transcripts most consistently downregulated in respiratory ciliated cells across twelve publicly available transcriptomic datasets from healthy airway donors. Transcript expression was analysed using Trailmaker™. Only transcripts with an $\log_2$ FC value of ≤ -1 were included. Colour intensity

represents FC $log_2$ in expression, with darker shades indicating stronger downregulation. White indicates absence of detectable expression in the corresponding donor sample.

## 3.2 Downstream analysis of candidate transcripts based on literature

To further assess the biological relevance of the transcripts identified through Trailmaker™, a literature review was conducted to evaluate their reported functions, cellular localisation, and potential involvement in viral entry processes. This approach allowed to prioritise the candidates that were not only enriched in ciliated cells but also had links to virus-host interactions (**Table 5**).

***Table 5. Summary of candidate transcripts enriched in respiratory ciliated cells.***
Transcripts were identified through transcriptomic analysis of respiratory ciliated cells. This table provides an overview of candidate transcript functions and localisation.

| Candidate transcript | Description | Localisation | Reference |
|---|---|---|---|
| Calmodulin 1 (CALM1) | Plays a key role in calcium signal transduction pathway, and it also has been found to have host-virus interactions, potentially contributing to the viral life cycle | Cytoplasm and plasma membrane | ('CALM1,' 2025) |
| Coiled-coil domain containing 78 (CCDC78) | Function includes cilium biogenesis and degradation | Cytoplasm and plasma membrane | ('CCDC78,' 2025) |

| | | | |
|---|---|---|---|
| Cadherin-related family member 3 (CDHR3) | Involved in cell adhesion and is a known host cell receptor for viral entry | Plasma membrane | ('CDHR3,' 2025) |
| Calcium and integrin binding 1 (CIB1) | Has a multitude of functions, some of them are apoptosis, cell adhesion, differentiation, etc | Cytoplasm and plasma membrane | ('CIB1,' 2025) |
| Dynein axonemal assembly factor 1 (DNAAF1) | Involved in the stability of ciliary architecture and regulation of cilia and microvilli | Cytoplasm and plasma membrane | ('DNAAF1,' 2025) |
| Dynein light chain LC8-type 1 (DYNLL1) | Involved in many processes, some of which are apoptosis, transcription, motile cilia assembly, etc. It is also involved in host-virus interactions through the mucociliary clearance, if this is impaired it leads to higher infection rates. | Cytoplasm and plasma membrane | ('DYNLL1,' 2025) |
| Ezrin (EZR) | Its role is to maintain the cell's shape, it may be involved in the major cytoskeletal structures | Plasma membrane | ('EZR,' 2025) |
| Heat shock protein 90 alpha family class A member 1 (HSP90AA1) | One of its roles is a molecular chaperone, which promotes structural maintenance, maturation, and the regulation of specific target proteins that might | Cytoplasm and plasma membrane | ('HSP90AA1,' 2025) |

| | | | |
|---|---|---|---|
| | be responsible for cell cycles or signal transduction. It is also involved in host-virus interactions as it is often proviral, for example, assisting viruses in stabilising their own proteins | | |
| Membrane spanning 4-domains A8 (MS4A8) | Encodes a member of the MS4A gene family with a unique expression pattern, enriched in differentiated cells especially in the airway. May be involved in signal transduction. | Plasma membrane | ('MS4A8,' 2025) |
| Presenilin enhancer, gamma-secretase subunit (PSENEN) | It is an essential subunit of the gamma-secretase complex, this complex is highly involved in the Notch and Wnt signalling pathways | Cytoplasm and plasma membrane | ('PSENEN,' 2025) |
| Transmembrane channel like 5 (TMC5) | Functions are yet to be determined, although it is thought to be a component of an ion channel | Cytoplasm and plasma membrane | ('TMC5,' 2025) |
| Tetraspanin 1 (TSPAN1) | It is a structural component of tetraspanin-enriched microdomains, thereby it is involved in functions such as | Cytoplasm and plasma membrane | ('TSPAN1,' 2025) |

| | | | |
|---|---|---|---|
| | adhesion, signal transduction, and protein trafficking | | |
| Zinc finger MYND-type containing 10 (ZMYND10) | Structure organisation and motility of the cilium | Plasma membrane | ('ZMYND10,' 2025) |

Three candidates emerged as the most promising candidates, these were CDHR3, CIB1, and EZR, with the rationale for their selection explained below. CDHR3 stood out due to its strong and consistent enrichment in respiratory ciliated cells (Basnet *et al.*, 2019), its established role as a viral receptor, and its function in cell adhesion, which may enhance cell-to-cell spread by infecting neighbouring cells ('CDHR3,' 2025).

CIB1 was selected based on evidence linking it to viral entry and replication in multiple viral systems. It has been shown to be essential for Kaposi's Sarcoma-Associated Herpesvirus (KHSV) entry, CIB1 knockdown was seen to reduce entry by 70% and overexpression increased entry by 60% (Bandyopadhyay *et al.*, 2014). Additionally, CIB1, alongside CIB2, has been identified as a host helper factor for HIV-1 replication and is required for optimal receptor-mediated viral entry (Godinho-Santos *et al.*, 2016). CIB1 enrichment in ciliated cells suggests a potential accessory role in facilitating viral entry in the respiratory epithelium.

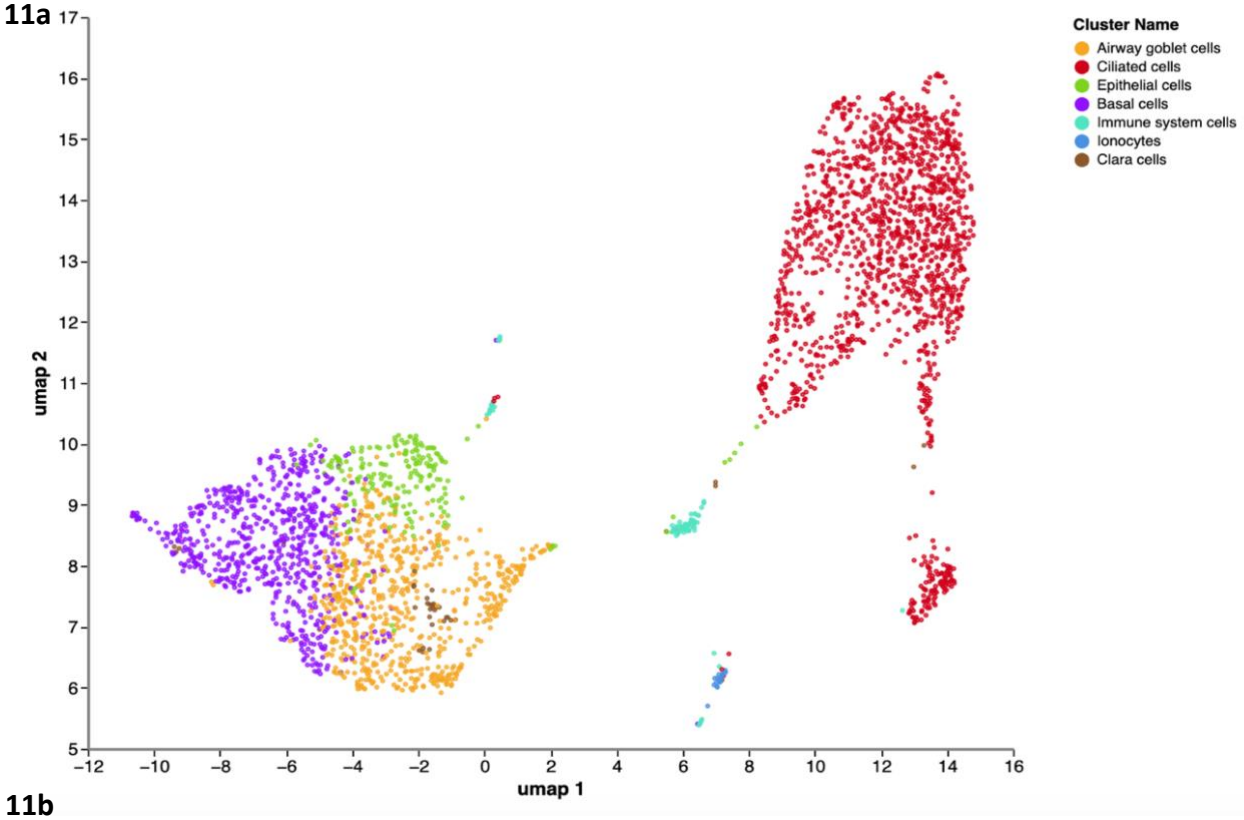
EZR was selected due to its established role in membrane dynamics and viral internalisation pathways. EZR is an important factor in the internalisation of the Japanese encephalitis virus, it forms a complex with caveolin-1 and Src kinase where it uses the caveolae-mediated entry pathway (Liu *et al.*, 2020). Importantly, EZR has also been reported to interact with the cytoplasmic tail of the SARS-CoV-1 spike protein and localised to viral

entry sites, with inhibition leading to reduced pseudovirus entry (Millet *et al.*, 2012). Although EZR exhibited broader expression across cell types compared to CDHR3, its functional relevance to viral entry justified its inclusion as a candidate.

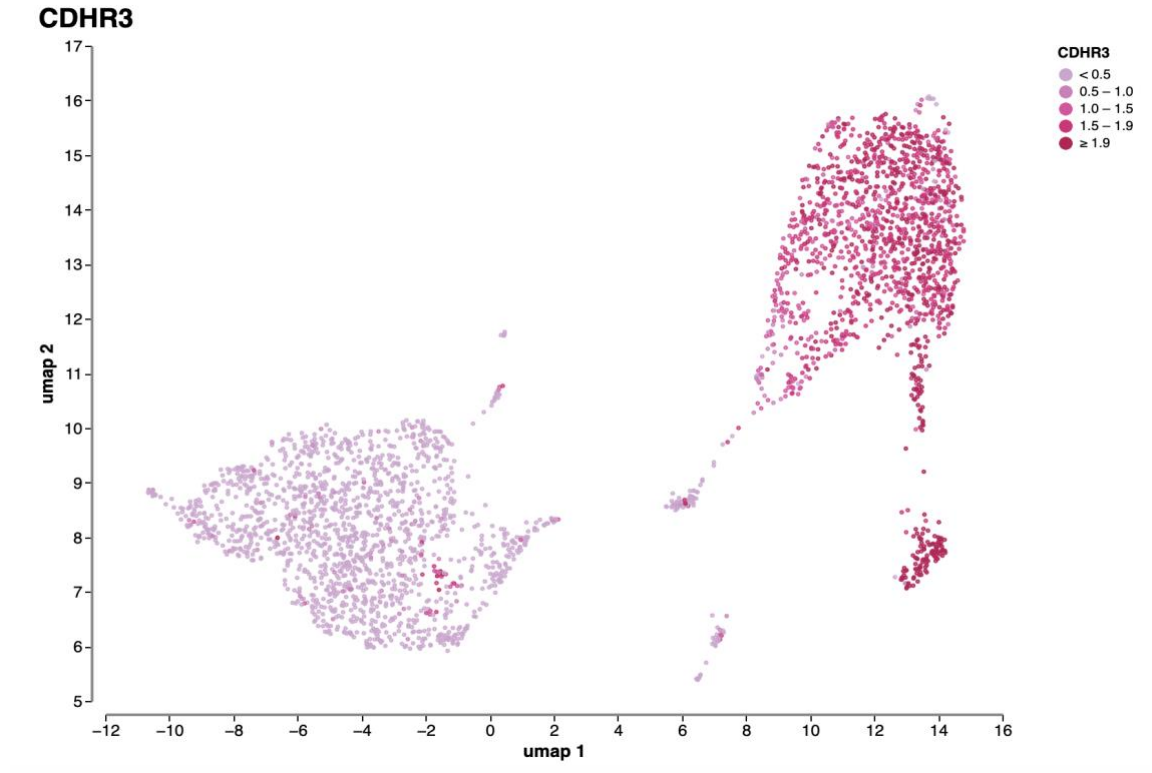
Several additional transcripts identified in the previous section were dismissed here due to limited evidence linking them to viral entry or because their known functions were primarily associated with ciliary structure, assembly, or maintenance. While these transcripts are highly relevant to ciliated cell biology, they are less likely to play a direct role in mediating viral attachment or entry. As a result, CDHR3, CIB1, and EZR were selected for further expression validation and downstream in silico and experiment analyses.

To confirm the enrichment of the three candidate transcripts in ciliated cells, their expression was visualised using Trailmaker™ (**Figs. 11**). **Figure 11a** shows an example of cell-type clustering from donor 6 (Paranjapye *et al.*, 2022), identifying goblet (orange), ciliated (red), epithelial (green), basal (purple), immune (light blue), ionocyte (dark blue), and clara (brown) cell populations. **Figures 11b-d** overlay CDHR3, CIB1, and EZR transcript expression onto the same UMAP, demonstrating strong enrichment in ciliated cells with lower-level expression in a limited number of other cell types, such as goblet cells.

**11a**


Cluster Name
Airway goblet cells
Ciliated cells
Epithelial cells
Basal cells
Immune system cells
Ionocytes
Clara cells
umap 2
umap 1

**11b**


CDHR3
CDHR3
< 0.5
0.5 – 1.0
1.0 – 1.5
1.5 – 1.9
≥ 1.9
umap 2
umap 1

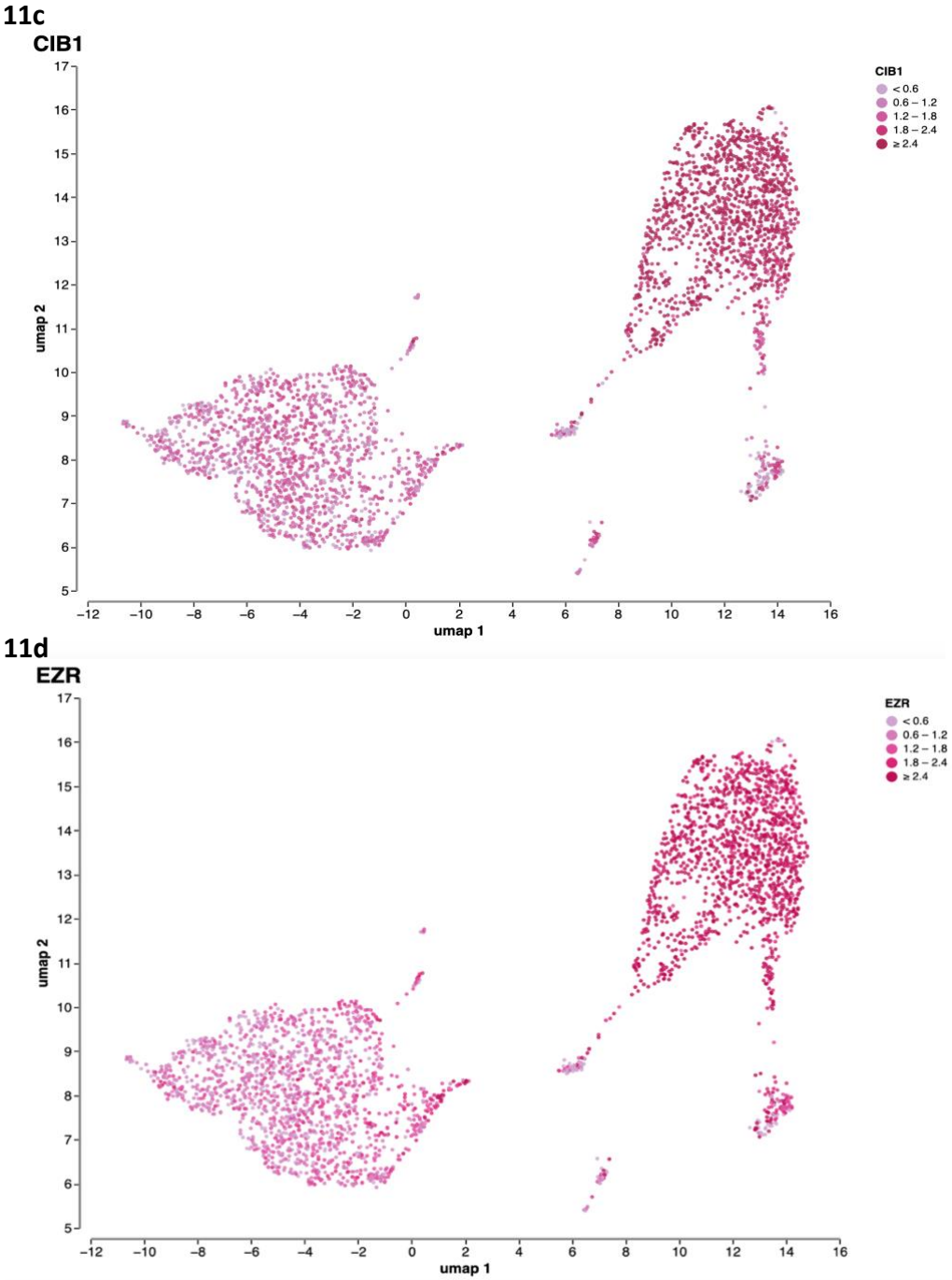



***Figure 10. Visualisation of CDHR3, CIB1, and EZR transcript expression in respiratory cell types.***

UMAP visualisation of single-cell transcriptomic data from a representative airway donor generated using Trailmaker™ (donor 6) (Paranjapye *et al.*, 2022). **(a)** Cell type clustering identifying different cell populations. **(b-d)** Overlay of CDHR3, CIB1, and EZR transcript expression, respectively, into the same UMAP. Higher expression is predominantly observed in ciliated cells, with lower expression detected in a limited number of other cell types.

## 3.3 Cross-dataset validation using The Human Protein Atlas and co-expression analysis

To further validate the candidate transcripts identified in Section 3.1, expression patterns were examined using The Human Protein Atlas, which provides tissue- and cell-type–specific expression information derived from multiple independent datasets. This analysis was used to assess whether the prioritised transcripts showed enrichment in respiratory ciliated cells beyond the datasets analysed in Trailmaker™.

In the ‘tissue-specific cell types’ category, CDHR3 showed the highest enrichment in respiratory ciliated cells, supporting its strong and consistent association with this cell type (**Fig. 12a**) ('CDHR3,' 2025). CIB1 also demonstrated enrichment in respiratory ciliated cells, although at a lower level and with broader expression across other tissues (**Fig. 12b**) ('CIB1,' 2025). Whereas *EZR* displayed comparatively low enrichment in respiratory ciliated cells (**Fig. 12c**) ('EZR,' 2025). These findings further support CDHR3 as the most ciliated cell specific among the selected candidates, while suggesting that CIB1 and EZR may play more generalised or accessory roles.

To examine patterns of co-expression across donor samples, a Jaccard similarity analysis was performed on the identified candidate transcripts. The resulting clustered heatmap (**Fig. 13**) reveals two distinct transcript modules. A core ciliary module, comprising *PSENEN, ZMYND10, DNAAF1, CCDC78, CALM1,* and *CDHR3*, exhibited high Jaccard similarity indices (0.80–0.95), indicating that these transcripts co-occur in nearly all donor samples and function together in ciliogenesis and ciliary maintenance. In contrast, an accessory/variable module consisting of IGFBP2, BASP1, AKAP9, and SPA17 showed lower similarity values (0.55–0.70), reflecting more variable or donor-specific expression.

Notably, CDHR3's tight clustering within the core ciliary module (Jaccard > 0.85 with its nearest neighbours) confirms its consistency and strength as a transcript of ciliated airway cells and underscores its potential role in ciliary adhesion and host–pathogen interactions.

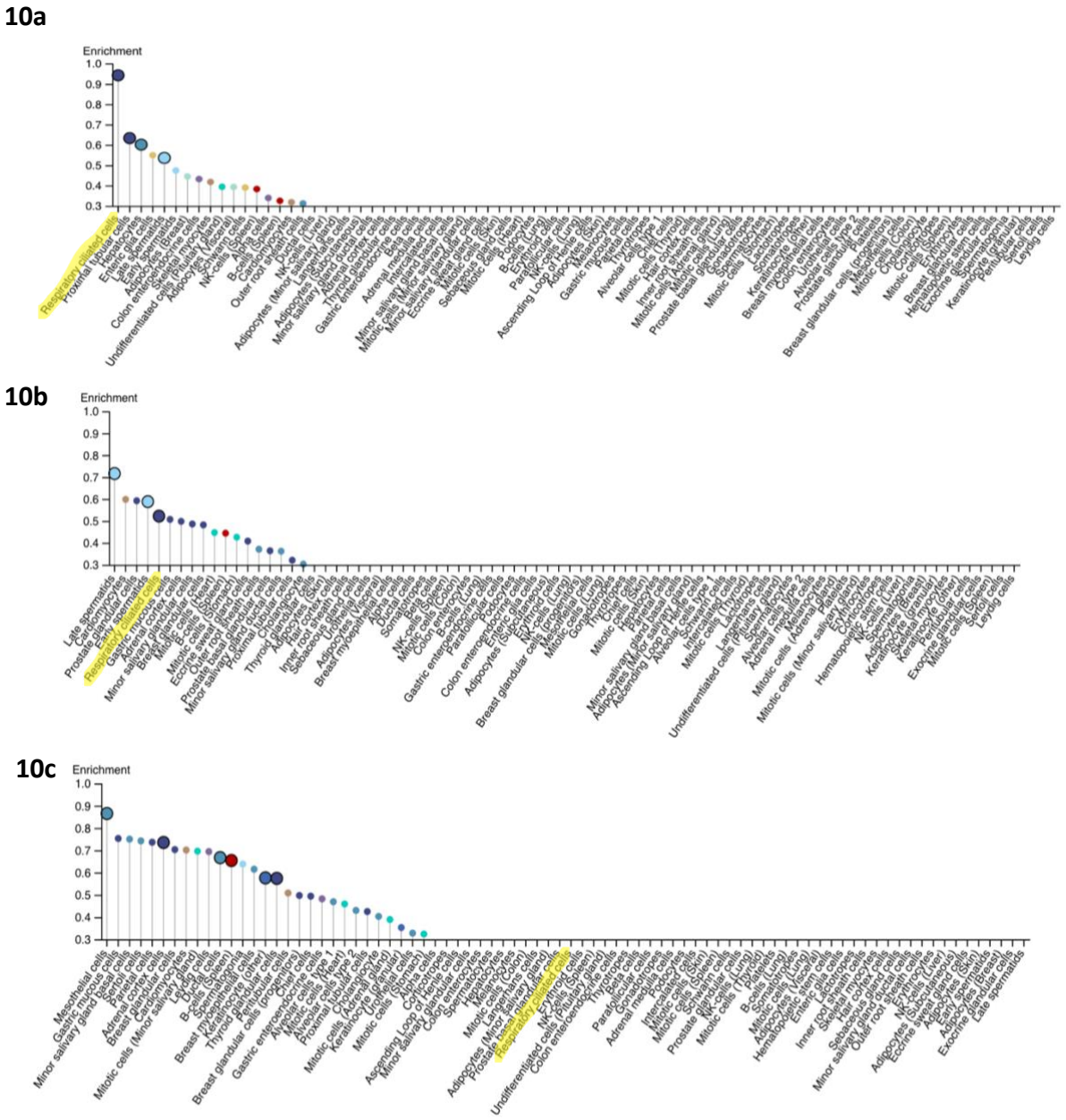



***Figure 11. Tissue-specific cell-type enrichment of CDHR3, CIB1, and EZR.***

**(a)** CDHR3, **(b)** CIB1, and **(c)** EZR transcripts, derived from The Human Atlas. Enrichment scores indicate relative transcript expression across profiled cell types, with higher values reflecting greater enrichment within a given cell type. Respiratory ciliated cells show the highest enrichment for CDHR3, moderate enrichment for CIB1, and low enrichment for EZR. Figures are adapted from The Human Protein Atlas ('CDHR3,' 2025; 'CIB1,' 2025; 'EZR,' 2025).

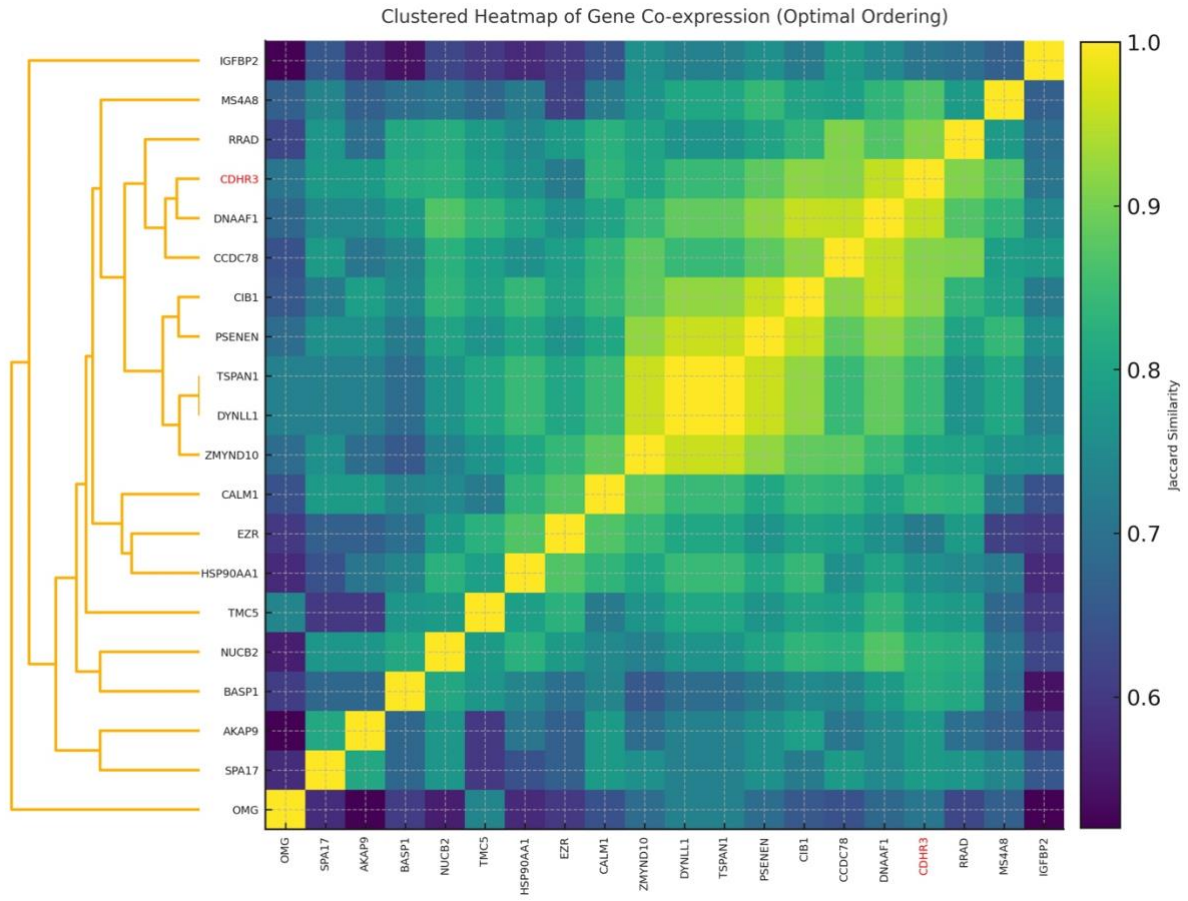



***Figure 12. Co-expression analysis of ciliated cell-associated transcripts across donor samples.***

The clustered heatmap displays pairwise Jaccard similarity indices for transcripts enriched in respiratory ciliated cells across donor samples. Jaccard indices reflect the proportion of shared expression between transcript pairs across datasets, with higher values indicating more consistent co-occurrence. Hierarchical clustering was applied to identify groups of co-expressed transcripts. A core ciliary module, including CDHR3 (highlighted in red), exhibits high similarity values, while a secondary module displays lower and more variable similarity across donors.

## 3.4 In silico docking of candidate host proteins with the HKU1 spike CTD

Molecular docking was performed to explore whether the candidate transcripts (CDHR3, CIB1, and EZR) could interact with the HKU1 spike protein at the structural level. Using protein-ligand docking, the primary objective is to predict potential binding modes

based on a known three-dimensional (3D) structure. Candidate dockings are then evaluated and ranked using a scoring function (Morris and Lim-Wilby, 2008).

The molecular structure of the HKU1 spike used in this study (PDB ID: 5I08) was obtained from the PDB, corresponding to a previously published cryo-electron microscopy structure (Kirchdoerfer *et al.*, 2016). The primary receptor binding domain in HKU1 is found in the CTD, therefore it is important to make sure this is used for the analysis (Wang *et al.*, 2024b). The HKU1 CTD sequence was identified using UniProt (Q5MQD0), visualised by PyMOL, and isolated following removal of water molecules and glycan residues (**Fig. 14**).

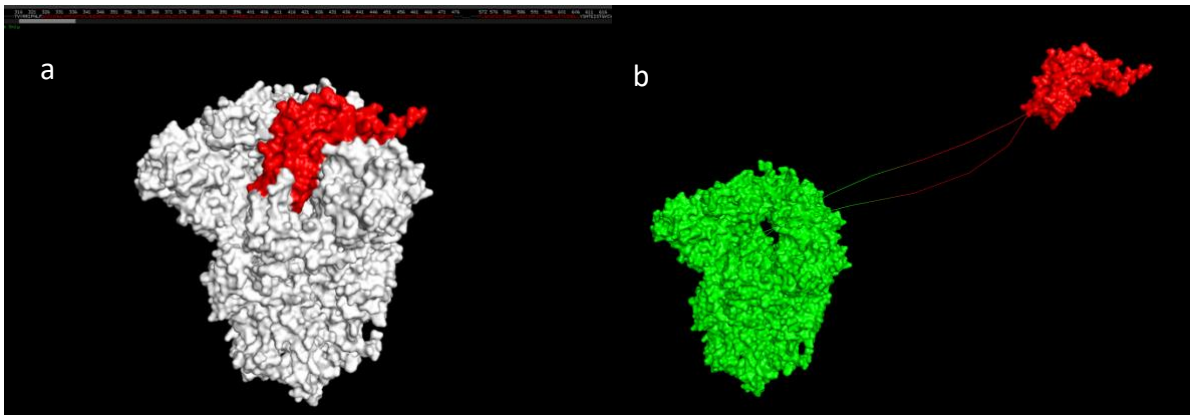



***Figure 13. Molecular structure of HKU1 spike protein, showing the highlighted and subsequently isolated CTD region.***

The RBD is located within the CTD, which is the focus of this analysis. This image was generated using PyMOL software, whilst the CTD sequence was obtained from UniProt.

Three-dimensional structures of the candidate transcripts (CDHR3, CIB1, and EZR) were obtained from the PDB and prepared using the same cleaning and visualisation workflow. The processed structures were then docked against the HKU1 CTD using the HDOCK server, which employs a hybrid docking strategy integrating template-based and template-free approaches (Yan *et al.*, 2020).

Docking simulations generated multiple predicted interaction models for each candidate protein. Visual inspection of the top-ranked docking poses indicated predicted contact interfaces between the HKU1 CTD and all three host proteins (**Fig. 15**). Analysis of

the docking output identified recurring spike residues involved in predicted interactions, with the five most frequent interacting residues for each candidate summarised in **Table 6**. Several spike residues were involved in multiple predicted contacts, suggesting preferential interaction regions within the CTD.

These docking results indicate that interactions between the HKU1 spike CTD and CDHR3, CIB1, and EZR are structurally possible. However, docking analyses are predictive and do not demonstrate direct or biologically relevant interactions *in vivo*. As such, these findings should be interpreted as hypothesis-generating rather than evidence of receptor interactions.

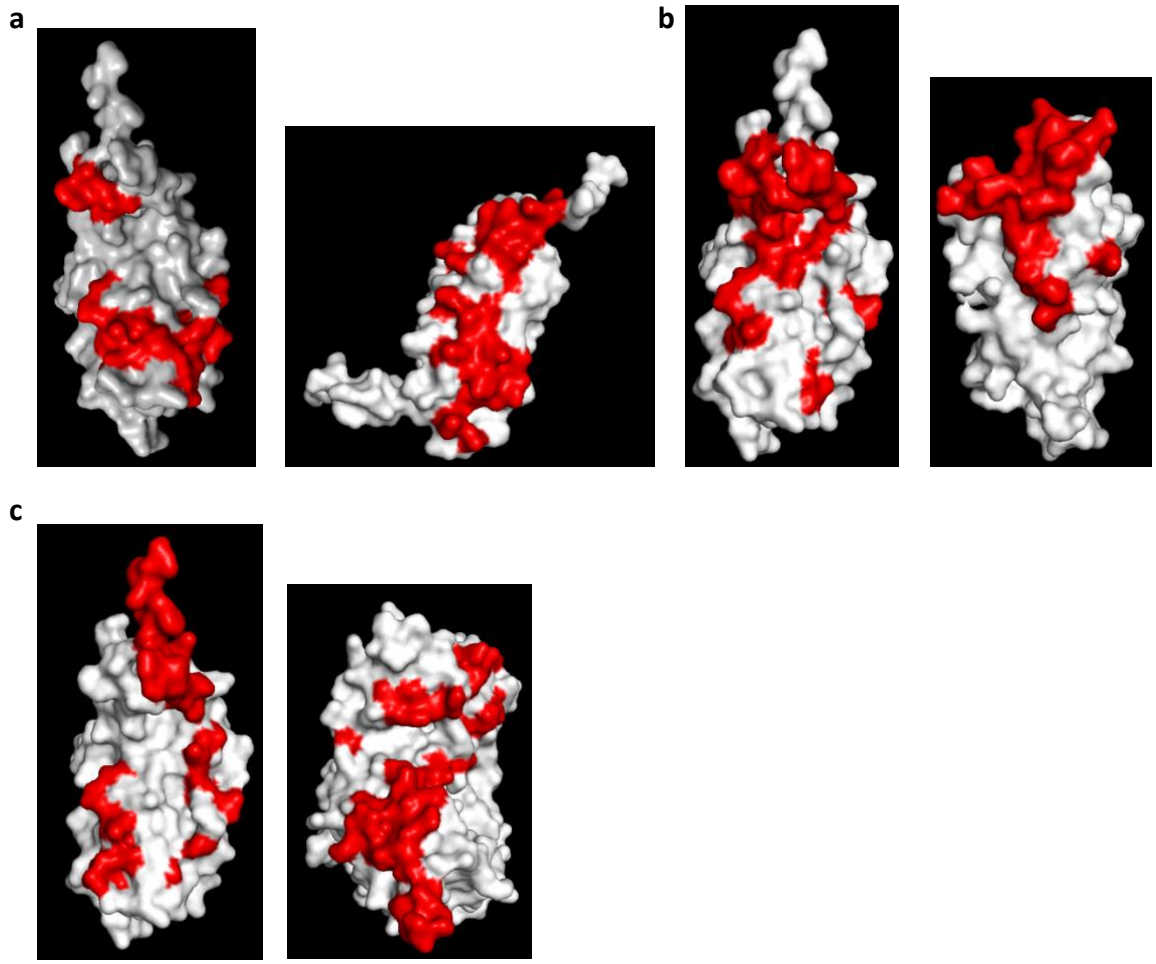



***Figure 14. Predicted docking interactions between HKU1 spike CTD and candidate transcripts.***

Representative docking poses generated by HDOCK showing predicted interactions. **(a)** HKU1 CTD (left) docked with CDHR3 (right). **(b)** HKU1 CTD (left) docked with CIB1 (right). **(c)** HKU1 CTD (left) docked with EZR (right). The red regions indicate predicted interaction areas between the HKU1 CTD and the respective transcript. Docking results represent *in silico* predictions and do not confirm biological interaction.

***Table 6. Predicted interacting residues between HKU1 spike CTD and candidate transcripts.***

Top five HKU1 spike residues predicted to interact with CDHR3, CIB1, and EZR based on HDCOK docking simulations. Docking was performed between the HKU1 spike CTD (isolate N5) (PDB: 5I08) (Kirchdoerfer *et al.*, 2016) and each candidate transcript. Residues are numbered according to chain A of the trimeric spike structure, used as a representative protomer. Interaction count indicates the frequency in which each spike residue was predicted to form contacts with the indicated transcript across docking models.

| | CIB1 | CDHR3 | EZR |
|---|---|---|---|
| Residues | Interaction Count | | |
| 32A | 10 | | |
| 43A | 6 | | |
| 45A | 7 | | |
| 122A | | 6 | |
| 124A | | 6 | |
| 151A | | | 7 |
| 326A | | 6 | |
| 348A | | 6 | |
| 415A | | | 6 |
| 442A | 8 | | |
| 444A | 6 | | |
| 472A | | | 6 |
| 473A | | | 9 |
| 476A | | | 9 |
| 595A | | 7 | |

## 3.5 Interpretation of transcriptomic findings

This chapter aimed to identify host factors enriched in respiratory ciliated cells that may contribute to viral entry, using multiple dataset transcriptomic analysis, literature review, and *in silico* docking. From 26 donor respiratory tract samples, a subset of consistently overexpressed transcripts was identified, from which CDHR3, CIB1, and EZR were prioritised based on ciliated-cell enrichment, predicted subcellular localisation, and prior evidence linking them to virus-host interactions.

Among the candidates, CDHR3 emerged as the strongest due to its high and specific expression in respiratory ciliated cells across donors, its tight clustering within a conserved ciliary gene module, and its established role as a viral receptor. These features suggest that CDHR3 may contribute to viral attachment at the ciliated cell surface. In contrast, CIB1 and EZR displayed broader expression patterns but were selected due to their known involvement in viral entry and intracellular processes. Their enrichment in ciliated cells indicates that they may function as accessory host factors that enhance or regulate viral entry rather than acting as primary receptors.

Molecular docking analysis predicted interactions between the HKU1 spike CTD and all three candidate proteins, supporting the structural possibility of these interactions. However, as docking approaches are predictive and do not account for cellular context or protein dynamics, these results should be interpreted as hypothesis-generating rather than confirmatory.

Overall, this analysis provides a rational basis for selecting CDHR3, CIB1, and EZR for downstream functional validation. Gain-of-function assays and pseudovirus entry experiments will be required to determine whether these transcripts directly mediate viral entry or act indirectly by influencing viral attachment or internalisation.

# Chapter 4: Optimisation of HCoV-HKU1 pseudovirus-based entry assays

This chapter describes the systematic development and optimisation of pseudovirus-based assays to investigate candidate transcript involvement in HCoV-HKU1 entry using gain-of-function approaches. Pseudoviruses provide a safe and experimentally tractable system for studying viral entry independently. HCoV-HKU1 was selected for investigation due to its resistance for cell culture propagation and the limited understanding of its cellular entry mechanisms. Therefore, pseudovirus systems represent a practical alternative for studying HKU1 entry while bypassing the technical constraints associated with live virus cultivation. A series of optimisation strategies were undertaken to establish an HKU1 pseudovirus entry assay, the approaches and outcomes are detailed in this chapter and provide important context for subsequent experimental strategies explored in the next chapter.

## 4.1 Assessment of producer cell lines for HKU1 pseudovirus production

HEK293T cells are commonly used for lentiviral pseudovirus production due to their high transfection efficiency (Millet and Whittaker, 2016). Initial HKU1 pseudovirus preparations were therefore generated using HEK293T cells following the standard production protocol described in section 2.4 (Temperton, Wright and Scott, 2015). However, the luciferase signals were consistently extremely low or undetectable, preventing reliable assessment of pseudovirus entry.

To determine whether producer cell type limited HKU1 pseudovirus production, HEK293T/27 cells were also evaluated. This cell line has been reported to perform more efficiently in transfection and transduction assays. HKU1 pseudoviruses were produced in parallel using HEK293T and HEK293T/27 cells under identical conditions.

Despite the use of two producer cell lines, no measurable improvement in HKU1 pseudovirus titre or detectible entry into target cells was observed for either condition. Therefore, it was not possible to identify producer cell choice as a limiting factor for HKU1 pseudovirus production or entry under the conditions tested.

## 4.2 Evaluation of HKU1 spike constructs for pseudovirus entry

Initial experiments utilised the full-length HKU1 spike glycoprotein to generate pseudoviruses, which were used to transduce a range of cell lines, including HEK293T, Caco-2, and Mv1Lu. At the time of these experiments, the role of TMPRSS2 in HKU1 entry had not yet been reported; therefore, early transductions were performed in cell lines that did not overexpress TMPRSS2. Following publication identifying TMPRSS2 as a key receptor for HKU1 (Saunders *et al.*, 2023), subsequent experiments were performed using HEK293T cells transiently overexpressing TMPRSS2. Despite these modifications, no detectable pseudovirus entry was observed in any of the tested cell lines when compared to positive and negative controls (**Fig. 16**).

To determine whether spike release from producer cells limited pseudovirus titres, an HKU1Del19 spike construct was subsequently evaluated. HKU1Del19 contains a 19-amino acid truncation at the C-terminus, removing part of the cytoplasmic tail. This modification has been reported to facilitate the release of pseudoviruses from packaging cells by preventing retention in the plasma membrane similar to SARS-CoV-2 PV (Chen *et al.*, 2021). Controls in all experiments included a negative control cell line transfected with an empty plasmid, which was transduced with a pseudovirus not containing any spike proteins. VSV-G protein is widely used for pseudotyping as it is pantropic, therefore it was used as the positive control pseudovirus (Hwang and Schaffer, 2013). However, pseudoviruses containing the

truncated HKU1Del19 spike also failed to produce measurable entry across all tested cell lines comparing to the negative and positive controls (**Fig. 17**).

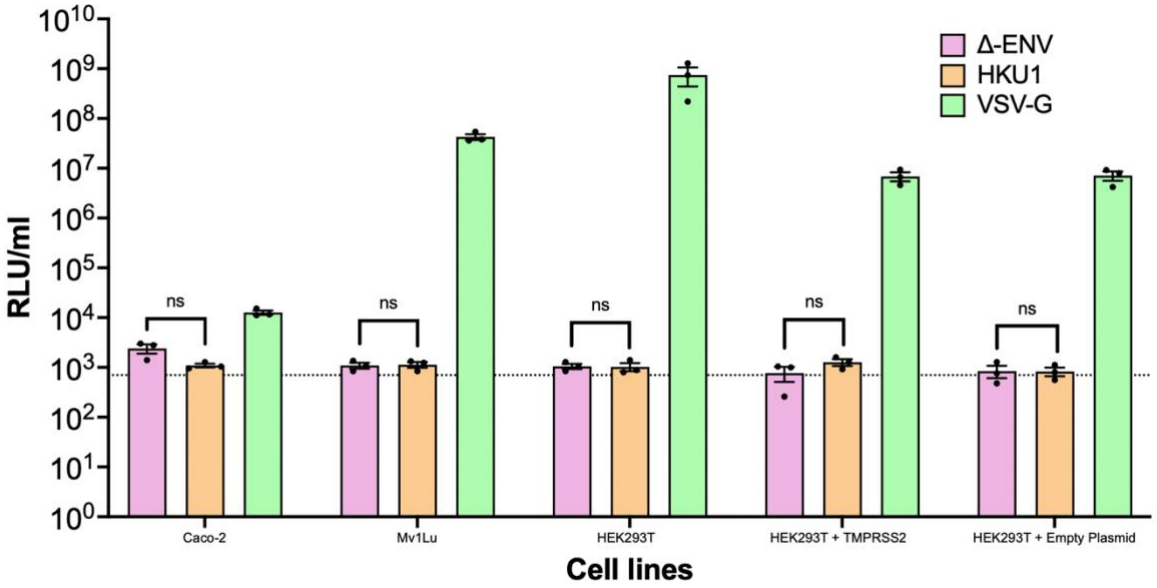



***Figure 15. Transduction of pseudoviruses containing full-length HKU1 spike glycoprotein into multiple cell lines.***

Cell lines included HEK293T, Caco-2, and Mv1Lu, as well as HEK293T cells transiently transfected to overexpress TMPRSS2, or an empty plasmid control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No detectable entry above background levels was observed in any of the tested cell lines. Data are presented as $\log_{10}$ RLU/ml.

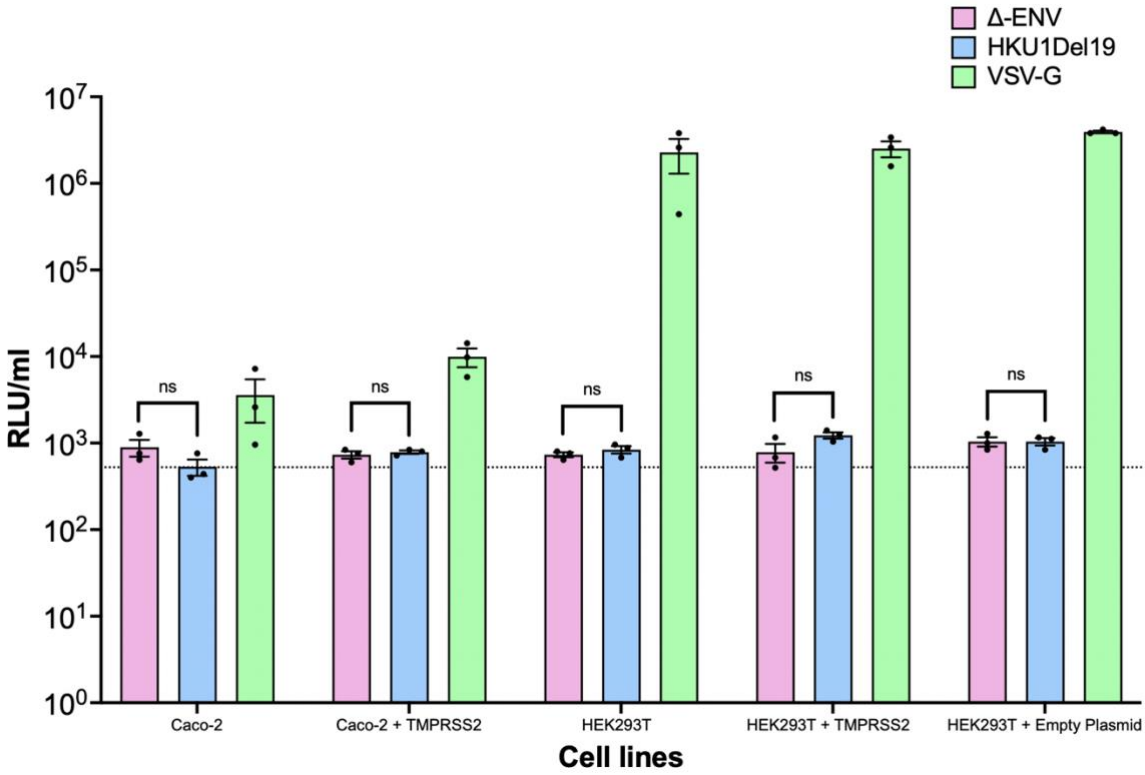



***Figure 16. Transduction of pseudoviruses containing truncated HKU1Del19 spike glycoprotein into Caco-2 and HEK293T cell lines.***

Cells were additionally transiently transfected to overexpress TMPRSS2 or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was measured using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No detectable entry above background levels was observed in any of the tested conditions. Data are presented as $\log_{10}$ RLU/ml.

## 4.3 Verification of reagents, plasmids, and transfection efficiency

To exclude the possibility that low HKU1 pseudovirus titres resulted from technical issues with reagents or plasmid quality, a series of quality control checks were performed. The transfection reagent in use was replaced with a fresh batch prior to subsequent

experiments. In addition, new plasmid preparations were generated from the original stocks to ensure optimal DNA concentration and integrity.

Following these updates, transfections were repeated using new reagents and plasmid preparations. Although the newly prepared plasmids yielded higher DNA concentrations, no improvement in HKU1 pseudovirus production or detectable entry was observed in transduction assays.

To directly assess transfection efficiency, a GFP-expressing plasmid was transfected into both HEK293T and HEK293T/27 producer cells. Live-cell imaging using the Incucyte system confirmed high levels of GFP expression in both cell lines, indicating efficient transfection (**Fig. 18**). These results demonstrate that transfection efficiency and reagent performance were not limiting factors in HKU1 pseudovirus production or entry under the conditions tested.

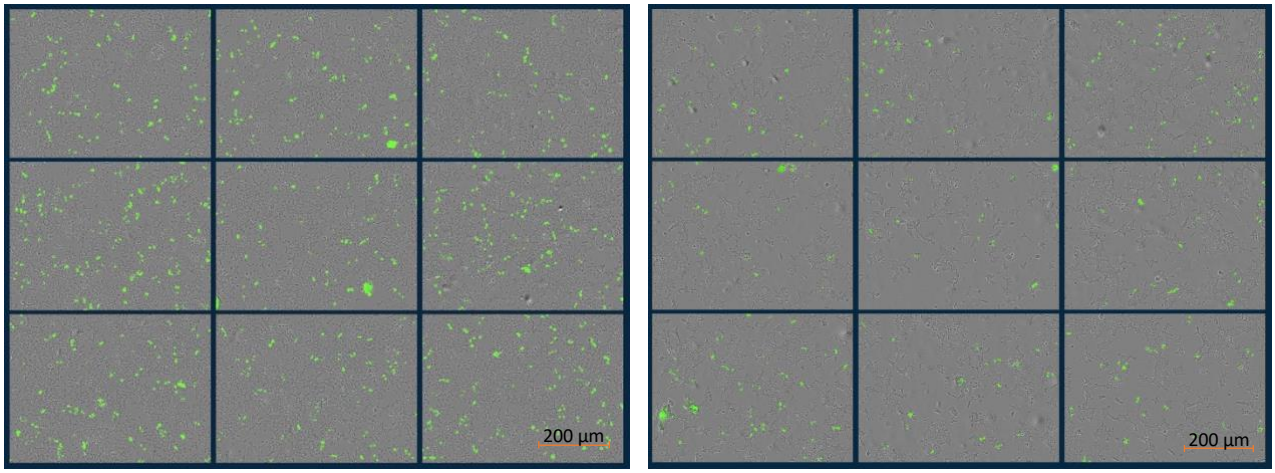



***Figure 17. HEK 293T (left) and HEK293T/27 (right) cells transfected with a GFP-expressing plasmid.***

GFP expression was visualised using Incucyte live-cell imaging system, confirming efficient transfection in both producer cell lines. Each image contains 9 sample points taken from a well in a 6-well plate.

## 4.4 Concentration of HKU1 pseudoviruses by PEG precipitation

To determine whether low HKU1 pseudovirus titres limited detectable entry, pseudovirus preparations were concentrated using a polyethylene glycol (PEG)-based virus precipitation method. The PEG Virus Precipitation Kit (Abcam, ab102538) was used on

HKU1Del19 pseudovirus preparations. The precipitated pseudoviruses were then used to transduce HEK293T and A549 target cells.

A549 cells continuously overexpressing ACE2 and TMPRSS2 were also included, alongside HEK293T cells transiently transfected to overexpress TMPRSS2, or an empty plasmid vector control. Despite concentration of pseudoviruses, no increase in luciferase signal relative to negative controls was detected in any of the tested cell lines (**Fig. 19**). These results indicate that insufficient pseudovirus concentration was unlikely to account for the absence of detectable HKU1 pseudovirus entry under the conditions tested.

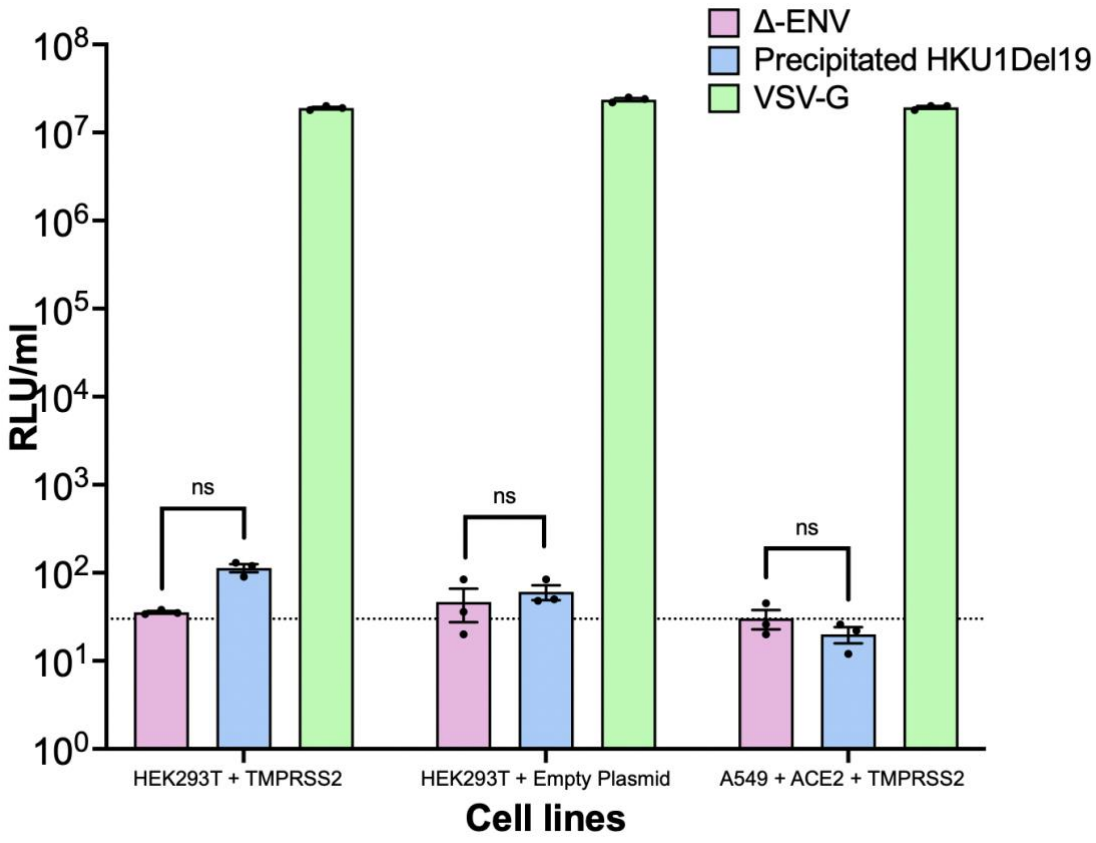



***Figure 18. Transduction of PEG-precipitated HKU1Del19 pseudoviruses into HEK293T and A549 target cell lines.***

A549 cells continuously overexpressing ACE2 and TMPRSS2, alongside HEK293T cells transiently transfected to overexpress TMPRSS2, or empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each

performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No increase in luciferase signal relative to negative controls was detected in any cell line. Data are presented as $\log_{10}$ RLU/ml.

## 4.5 Optimisation of pseudovirus harvest timing

To assess whether the timing of pseudovirus collection influenced detectable HKU1 entry, the harvest schedule was modified. Standard collections performed at 48 hours and 72 hours post-transfection were replaced with a single collection at 48 hours to minimise potential dilution from later, lower-titre supernatants.

HKU1 pseudoviruses produced under the modified collection schedule were used to transduce target cell lines under the same conditions as previous experiments. However, no increase in luciferase signal relative to negative controls was observed following the altered harvest timing, indicating that collection schedule did not account for the lack of detectable HKU1 pseudovirus entry.

## 4.6 Role of spike processing: neuraminidase and TMPRSS2

Proteolytic processing of coronavirus spike proteins is required for activation and subsequent viral entry (Hoffmann *et al.*, 2020). In standard pseudovirus production protocols, neuraminidase is commonly included to facilitate viral release from producer cells by cleaving sialic acids on the cell surface (Temperton *et al.*, 2007). In addition, host proteases such as TMPRSS2 have been shown to play key roles in spike priming for several coronaviruses, including HKU1 (Bertram *et al.*, 2013; Saunders *et al.*, 2023).

To assess whether spike processing limited HKU1 pseudovirus entry, pseudoviruses were produced under multiple conditions incorporating neuraminidase and TMPRSS2. These included production in the absence of both factors; neuraminidase alone (final concentration

in culture 0.1U/ml); TMPRSS2 alone; combined neuraminidase (final concentration in culture 0.1U/ml) and TMPRSS2; and neuraminidase at an increased final concentration (0.2U/ml) in combination with TMPRSS2. Resulting pseudoviruses were used to transduce HEK293T target cells transiently overexpressing TMPRSS2, or an empty plasmid vector control. Across all conditions tested, no statistical significance in luciferase signal relative to negative controls was detected (**Fig. 20**).

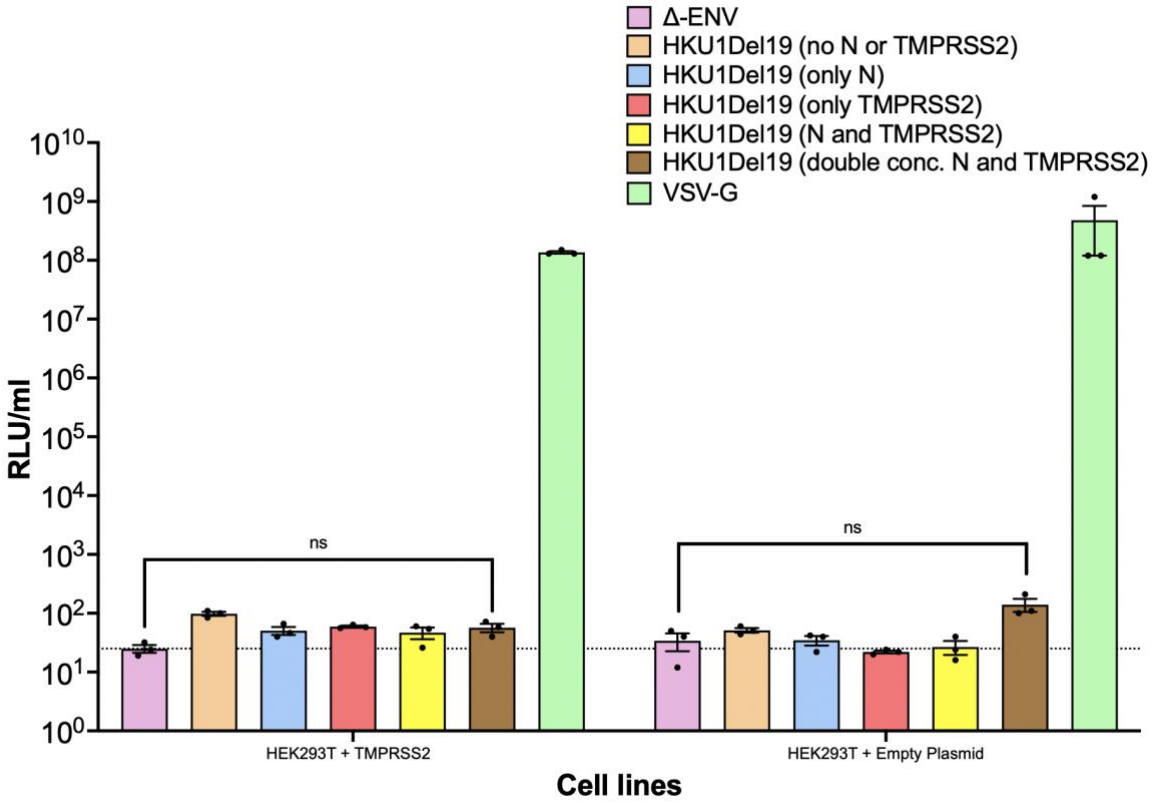



***Figure 19. Effect of neuraminidase and TMPRSS2 on HKU1 pseudovirus entry.***

PV transduction of target cells using HKU1 pseudoviruses produced under varying neuraminidase (N) and TMPRSS2 conditions. Pseudoviruses were generated in the absence of both factors; with neuraminidase alone (final concentration in culture 0.1 U/mL); TMPRSS2 alone; combined neuraminidase (final concentration in culture 0.1 U/mL) and TMPRSS2; or neuraminidase at an increased final concentration (0.2 U/mL) in combination with TMPRSS2. HEK293T target cells were transiently transfected to overexpress TMPRSS2, or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase

reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. No increase in luciferase signal relative to negative controls was detected under any condition. Data are presented as $\log_{10}$ RLU/ml.

## 4.7 Evaluation of target cell lines for HKU1 pseudovirus entry

To assess whether target cell permissiveness limited detectable HKU1 pseudovirus entry, a range of cell lines was evaluated. HEK293T cells were used as the primary target cell line throughout optimisation due to their robust growth and transfection efficiency. In addition, Chinese hamster ovary (CHO) cells were tested following reports that HKU1 pseudoviruses may show improved entry in this cell type.

CHO cells were transiently transfected to overexpress TMPRSS2, or an empty plasmid vector control prior to transduction with HKU1Del19 pseudoviruses. No increase in luciferase signal relative to negative controls was detected under any condition (**Fig. 21**).

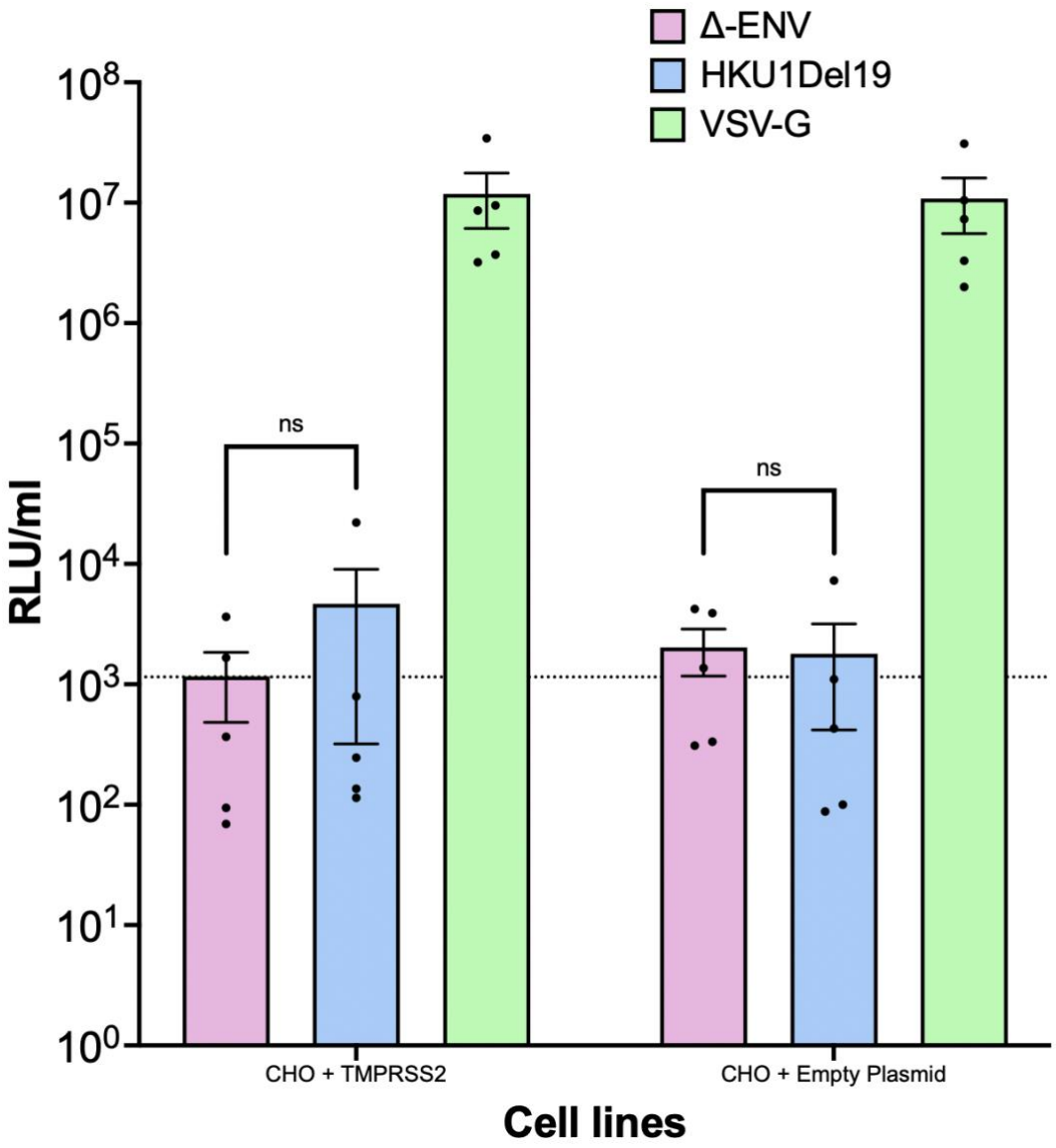



***Figure 20. Evaluation of CHO cells as target cells for HKU1Del19 pseudovirus entry.***

Transduction of CHO cells with HKU1Del19 pseudoviruses following transient transfection with TMPRSS2, or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 5 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No increase in luciferase signal relative to negative controls was detected under any condition. Data are presented as $\log_{10}$ RLU/ml.

A549 cells continuously overexpressing ACE2 and TMPRSS2 were also used as target cells, due to their widespread use in coronavirus entry studies, to assess whether HKU1Del19 pseudoviruses could enter. Despite their use in previous research, no significant detectable increase in pseudovirus entry was observed (**Fig. 22**).

Together, these results indicate that variation in target cell line and transcript expression was insufficient to enable the detection of HKU1 pseudovirus entry under the conditions tested.

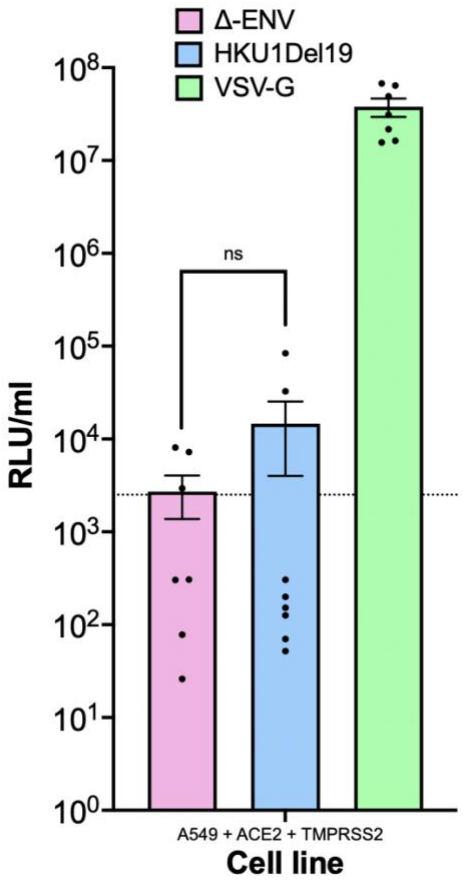



***Figure 21. Evaluation of A549 cells as target cells for HKU1 pseudovirus entry.***

Transduction of A549 cells continuously overexpressing ACE2 and TMPRSS2 with HKU1Del19 pseudoviruses. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 8 independent experiments, each performed with technical duplicates. Statistical significance was determined using one-way ANOVA with Dunnett's multiple comparison test. *p ≤ 0.05, **p ≤ 0.01, ***p ≤

0.001. No detectable increase in luciferase signal relative to negative control was observed. Data are presented as $\log_{10}$ RLU/ml.

## 4.8 Inclusion of additional HKU1 structural proteins (M, E, and HE)

To assess whether inclusion of additional structural components could enhance pseudovirus assembly or entry, pseudoviruses were generated containing the membrane (M), envelope (E), and hemagglutinin esterase (HE) proteins alongside the HKU1Del19 spike. Inclusion of these proteins was intended to represent the wild type virion more closely and potentially improve viral attachment and entry.

Pseudoviruses containing HKU1Del19 spike in combination with M, E, and HE were used to transduce A549 cells continuously overexpressing ACE2 and TMPRSS2. However, inclusion of M, E, and HE did not result in a significant increase in pseudovirus entry relative to negative control (**Fig. 23**).

These findings indicate that the addition of HKU1 structural proteins M, E, and HE was insufficient to enable statistically significant HKU1 pseudovirus entry under the conditions tested.

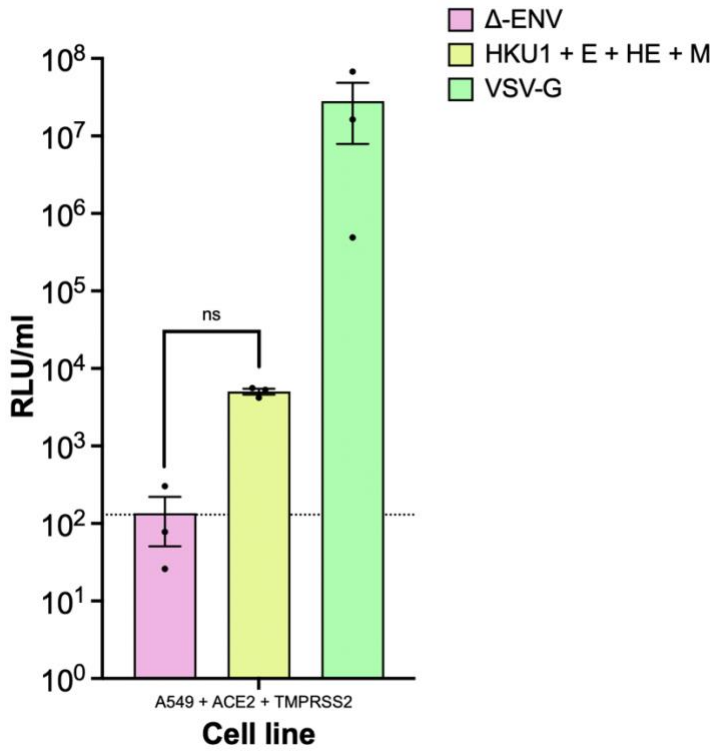



***Figure 22. Effect of additional HKU1 structural proteins on pseudovirus entry.***

Transduction of A549 cells continuously overexpressing ACE2 and TMPRSS2 with HKU1Del19 pseudoviruses produced in the presence of additional HKU1 structural proteins (M, E, and HE). VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using one-way ANOVA with Dunnett's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No statistically significant increase in luciferase signal relative to negative control was observed following inclusion of M, E, and HE. Data are presented as $\log_{10}$ RLU/ml.

## 4.9 Interpretation of HKU1 optimisation experiments

This chapter aimed to establish a lentiviral pseudovirus-based system to investigate HCoV-HKU1 entry and enable functional analysis testing of candidate transcripts identified from prior transcriptomic analyses. However, despite extensive optimisation, reproducible HKU1 pseudovirus entry was not detected under any of the conditions tested.

A broad range of optimisations was explored, including spike construct design, producer, and target cell lines, pseudovirus concentration, harvest timing, spike processing, and inclusion of additional viral structural proteins. The consistent absence of detectable entry across all conditions indicates that the observed limitation is unlikely to arise from a single technical factor, but instead reflects inherent challenges associated with modelling HKU1 entry using lentiviral pseudovirus systems.

In contrast to other coronaviruses such as SARS-CoV-1 and SARS-CoV-2, whose spike proteins readily mediate pseudovirus entry, the HKU1 spike appears less compliant to lentiviral pseudotyping. Modifications, including cytoplasmic tail truncation, protease supplementation, and receptor overexpression, were insufficient to enable HKU1 entry in this study. This indicates that HKU1 entry may depend on additional host factors or specific cellular conditions that are not adequately reproduced in standard *in vitro* systems.

Overall, these findings highlight the technical challenges associated with studying HKU1 entry using pseudovirus models and provide important context for the experimental strategy implemented in the next chapter, where further investigation was performed to other respiratory viruses with established pseudovirus entry systems, enabling functional evaluation of candidate transcripts, as described in Chapter 5.

# Chapter 5: Evaluation of candidate transcripts in the entry of diverse respiratory viruses

Due to the limited ability to assess the transcript involvement in HKU1 pseudovirus entry assay alone, this chapter expands the investigation to include a broader panel of respiratory viruses. These include additional coronaviruses (MERS-CoV, SARS-CoV-1, SARS-CoV-2 wild-type and its variants Alpha, Beta, Delta, BA.1, BA.2, BA.4, BA.5, XBB, XQ1.1), enabling comparison of the transcripts' effect across the coronavirus family while capturing the evolutionary diversity within SARS-CoV-2. Influenza type A viral isolates (H5N10) were also included to determine whether any observed effects extend beyond coronaviruses. HKU1 was retained within this panel to allow continued evaluation of candidate transcripts in the context of this virus, alongside systems in which viral entry could be more reliably quantified. By examining multiple viruses in parallel, this approach enables assessment of whether candidate transcripts influence viral entry in a virus-specific manner or contribute to more generalised mechanisms of respiratory virus entry. This comparative framework therefore provides a stronger basis for evaluating the functional relevance of candidate transcripts identified in earlier analyses.

## 5.1 Validation of pseudovirus entry systems

Pseudoviruses expressing glycoproteins of SARS-CoV-1, SARS-CoV-2, MERS-CoV, and H5N10 (Influenza A) were generated and assessed for their ability to enter permissive target cell lines. Strong numerical luciferase signals were observed for SARS-CoV-1 and SARS-CoV-2, confirming successful pseudovirus production and entry in ACE2-expressing cells (**Fig. 24**). In contrast, MERS-CoV pseudovirus entry was not detected, consistent with the absence of the required DPP4 receptor in the target cell lines. As well as H5N10

pseudovirus entry produced lower signals than SARS across the experiments. These results demonstrate that functional pseudovirus entry systems were established for selected viruses, enabling subsequent evaluation of candidate transcripts.

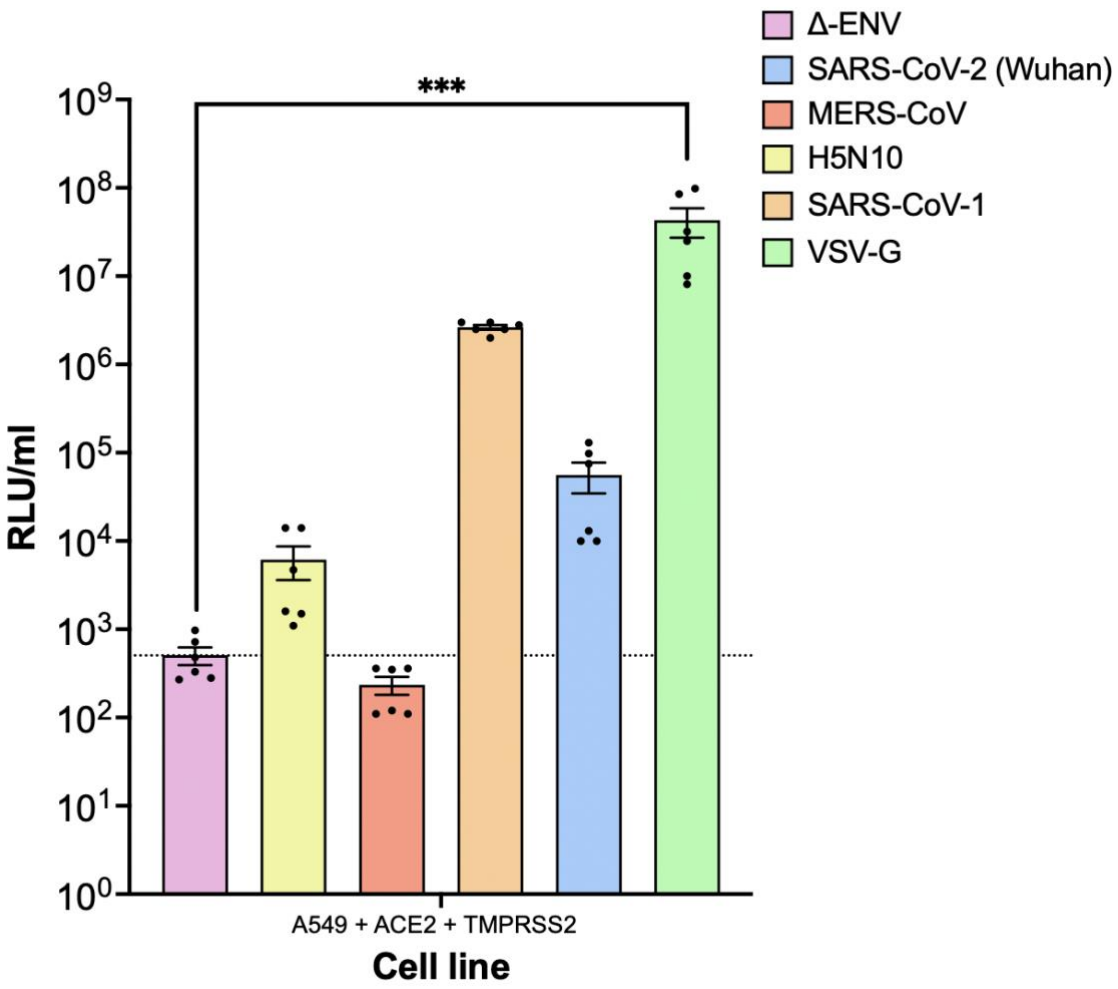



***Figure 23. Validation of pseudovirus entry across multiple respiratory viruses.***

Target cell lines were transduced with pseudoviruses representing SARS-CoV-1, SARS-CoV-2, MERS-CoV, and H5N10 (influenza A). VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Statistically significant entry was observed only in the positive control; however, SARS-CoV-1 exhibited the highest numerical entry values among the pseudoviruses, followed closely by SARS-CoV-2. Pseudovirus entry was assessed using a luciferase reporter assay. Data are presented as mean ± SEM from n = 6 independent experiments, each performed with technical duplicates. Data are presented as $\log_{10}$ RLU/ml.

## 5.2 Functional evaluation of candidate transcripts in HCoV-HKU1 pseudovirus entry

This section evaluates the role of candidate transcripts CDHR3, CIB1 and EZR in HKU1 pseudovirus entry model systems. Due to the inability to reliably detect HKU1 pseudovirus entry in earlier optimisation experiments, it remains unclear whether this reflects inefficient pseudovirus production, limitations in target cell permissiveness, or a combination of both. Nevertheless, these candidate transcripts were assessed to determine whether their expression could influence detectable HKU1 pseudovirus entry under the conditions tested.

HKU1Del19 pseudoviruses were used to transduce CHO cells and A549 cells continuously overexpressing ACE2 and TMPRSS2. CHO cells were transiently transfected to overexpress CDHR3, CIB1, EZR, TMPRSS2, CDHR3 + TMPRSS2 or an empty plasmid vector control prior to transduction. No significant increase in luciferase signal relative to negative controls was observed in most cell lines (**Fig. 25**). However, in one experimental replicate, CHO cells overexpressing CDHR3 alone and in combination with TMPRSS2 exhibited an increased luciferase signal, suggesting a potential enhancement of viral entry. This effect could not be observed in subsequent independent experiments and was therefore not reproducible, indicating that the initial observation may reflect experimental variability rather than a consistent biological effect.

Similarly, A549 cells continuously overexpressing ACE2 and TMPRSS2 were transiently transfected to overexpress CDHR3, CIB1, EZR, or an empty plasmid vector control prior to transduction. No significant HKU1 pseudovirus entry was observed (**Fig. 26**).

Overall, these results indicate that overexpression of CDHR3, CIB1, EZR, or TMPRSS2 did not result in significant increase of HKU1 pseudovirus entry in the model systems used. However, given the lack of measurable baseline entry, it is not possible to

determine whether these transcripts have a functional role in HKU1 entry under these experimental conditions.

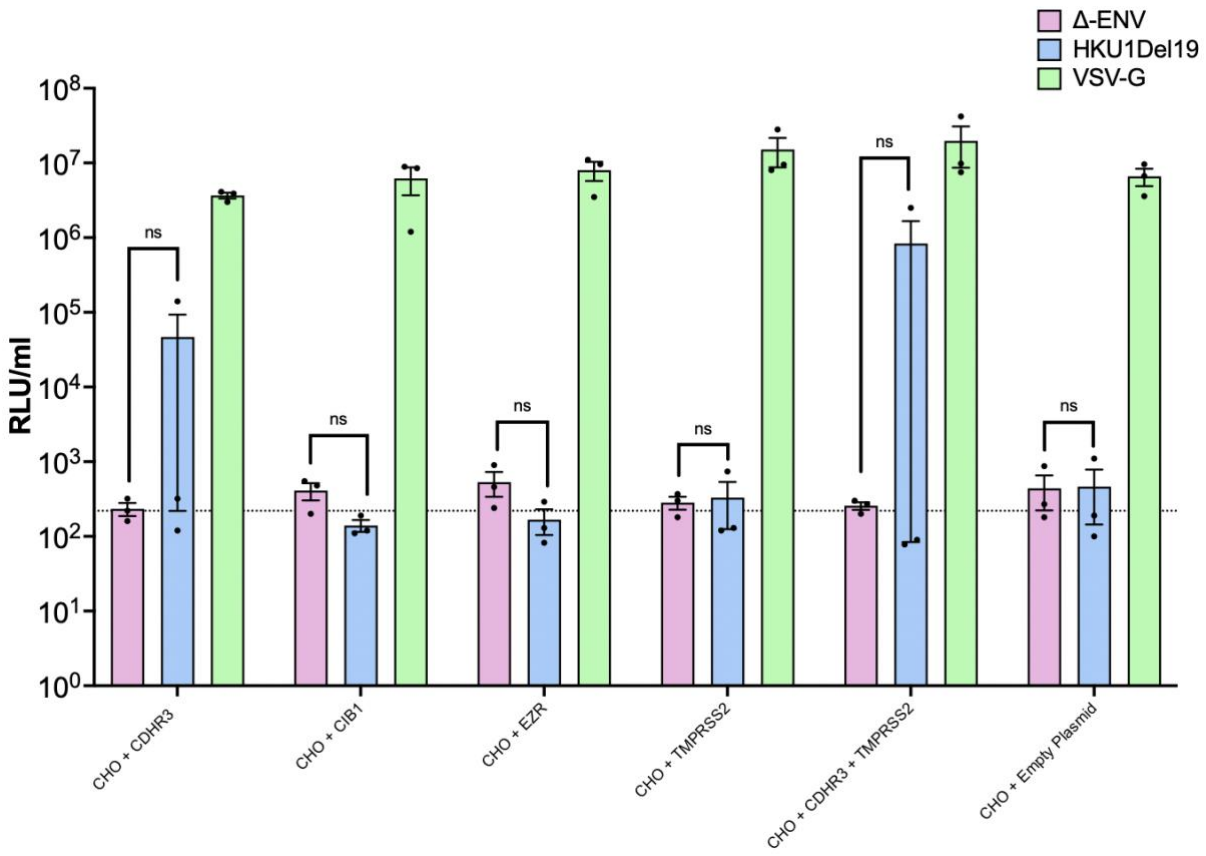



***Figure 24. Effect of CDHR3, CIB1, EZR, and TMPRSS2 overexpression on HKU1Del19 pseudovirus entry in CHO cells.***

CHO cells were transduced with HKU1Del19 pseudoviruses following transient transfection with CDHR3, CIB1, EZR, TMPRSS2 or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No significant increase in luciferase signal relative to negative controls was observed. However, in one experimental replicate, CHO cells overexpressing CDHR3 alone and in combination with TMPRSS2 exhibited an increased luciferase signal, suggesting a potential enhancement of viral entry. This effect could not be observed in subsequent independent experiments and was therefore not reproducible. Data are presented as $\log_{10}$ RLU/ml.

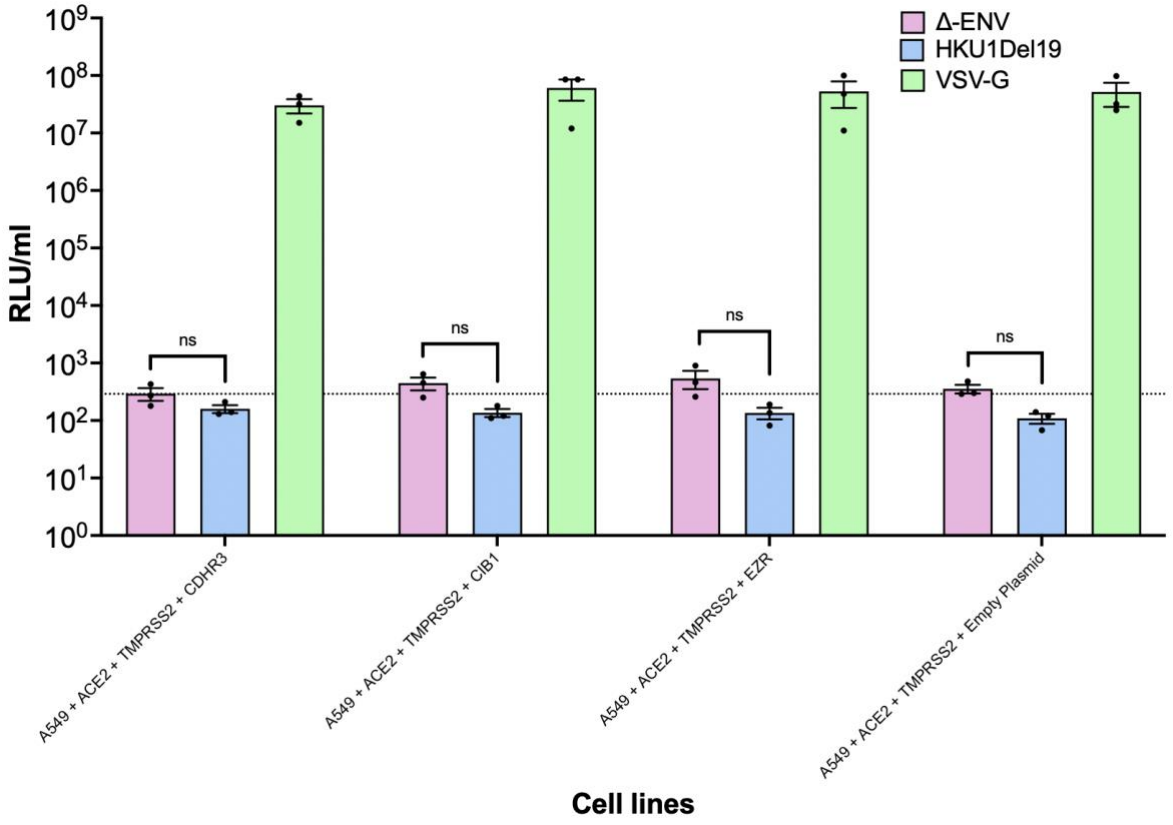



***Figure 25. Effect of CDHR3, CIB1, or EZR overexpression on HKU1Del19 pseudovirus entry in A549 cells.*** A549 cells continuously expressing ACE2 and TMPRSS2 transduced with HKU1Del19 pseudoviruses following transient transfection with CDHR3, CIB1, EZR or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. No significant increase in luciferase signal relative to negative controls was observed. Data are presented as $\log_{10}$ RLU/ml.

## 5.3 Functional evaluation of candidate transcripts in SARS-CoV-1 pseudovirus entry

To assess whether candidate transcripts influence SARS-CoV-1 entry, A549 cells continuously overexpressing ACE2 and TMPRSS2 were transiently transfected with CDHR3, CIB1, EZR, or an empty plasmid vector control prior to pseudovirus transduction. SARS-

CoV-1 pseudovirus entry was detectable across all conditions, with luciferase signals noticeably higher than the negative control (**Fig. 27**). However, no statistically significant differences were observed between transcript overexpression conditions and the control. Although increased numerical values were observed relative to the negative control, these were consistent across all conditions, indicating no transcript-specific enhancement of viral entry.

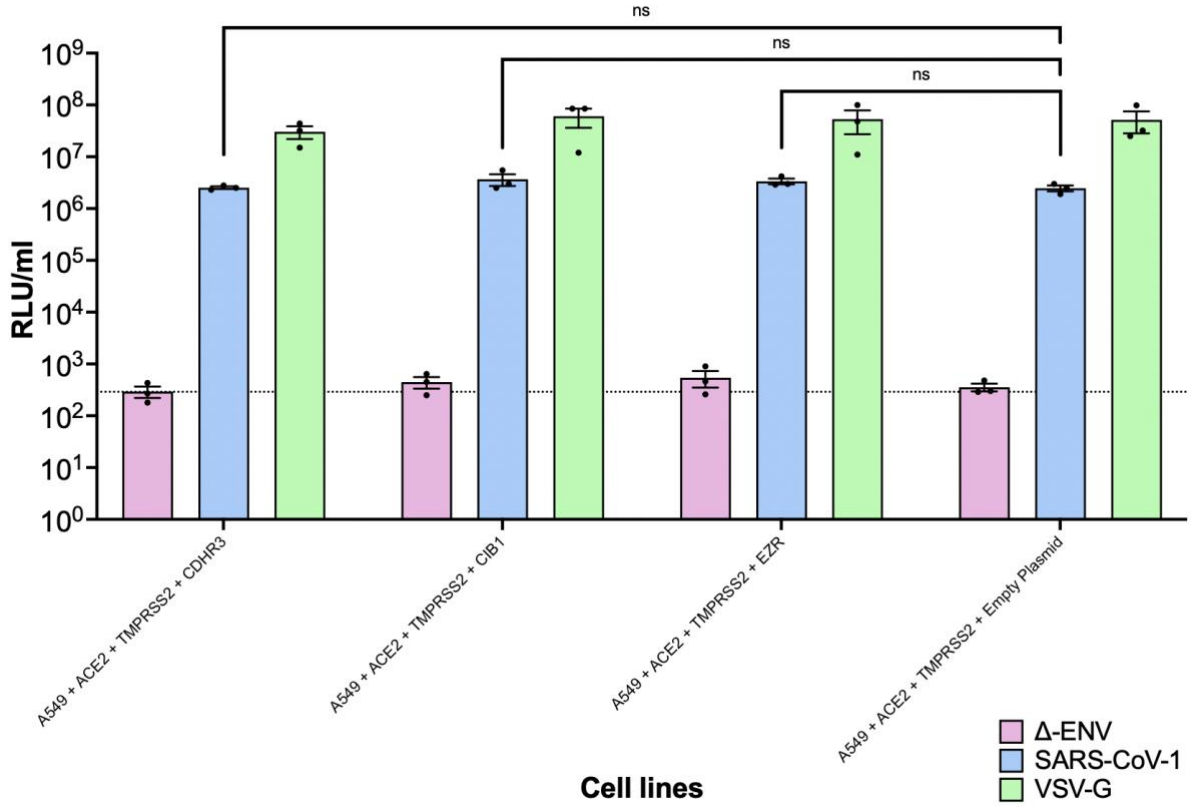



***Figure 26. Effect of CDHR3, CIB1, or EZR overexpression on SARS-CoV-1 pseudovirus entry in A549 cells.*** A549 cells continuously expressing ACE2 and TMPRSS2 transduced with SARS-CoV-1 pseudoviruses following transient transfection with CDHR3, CIB1, EZR or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No statistically significant differences in luciferase signal were observed between conditions relative to the negative control. Although higher numerical values were observed across all conditions compared to the

negative control, these differences were consistent across transcript overexpression conditions and therefore do not indicate a transcript-specific effect on SARS-CoV-1 pseudovirus entry. Data are presented as $\log_{10}$ RLU/ml.

## 5.4 Functional evaluation of candidate transcripts in MERS-CoV pseudovirus entry

To assess whether candidate transcripts influence MERS-CoV entry, A549 cells continuously overexpressing ACE2 and TMPRSS2 were transiently transfected with CDHR3, CIB1, EZR, or an empty plasmid vector control prior to pseudovirus transduction. No detectable increase in luciferase signal was observed across any condition (**Fig. 28**). This is consistent with the absence of the primary MERS-CoV entry receptor, dipeptidyl peptidase-4 (DPP4), in the target cell lines used, indicating that the system was not permissive for MERS-CoV entry under the conditions tested (Guarner, 2020).

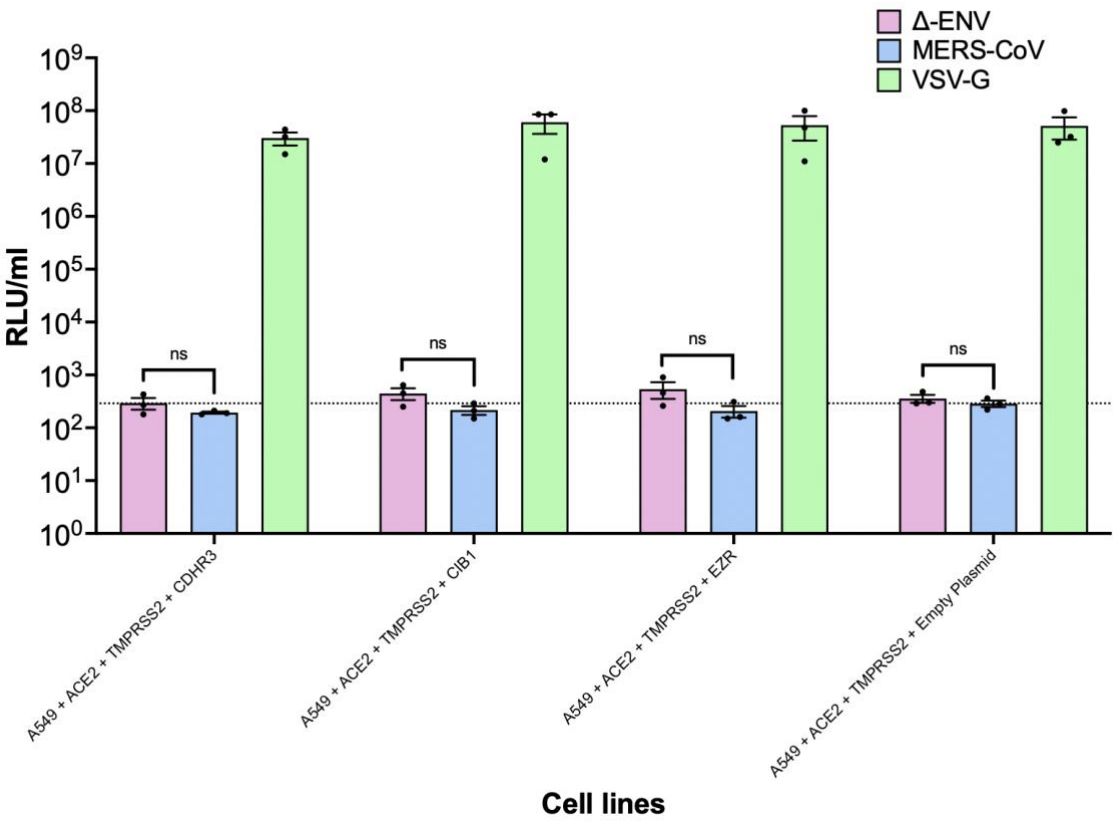



***Figure 27. Effect of CDHR3, CIB1, or EZR overexpression on MERS-CoV pseudovirus entry in A549 cells.***

A549 cells continuously expressing ACE2 and TMPRSS2 transduced with MERS-CoV pseudoviruses following transient transfection with CDHR3, CIB1, EZR or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No significant increase in luciferase signal relative to negative controls was observed. Data are presented as $\log_{10}$ RLU/ml.

## 5.5 Functional evaluation of candidate transcripts in influenza type A pseudovirus entry

Influenza isolate H5N10 pseudovirus entry was assessed to evaluate whether candidate transcripts influence entry of viruses utilising alternative receptor mechanisms. Low levels of luciferase signal were observed across all conditions, with no statistically significant differences detected between transcript overexpression and control conditions (**Fig. 29**). This may reflect a lack of appropriate sialic acid receptors in the cell lines used (Dou *et al.*, 2018). Overall, these findings suggest that overexpression of CDHR3, CIB1, or EZR does not enhance H5N10 pseudovirus entry in this system.

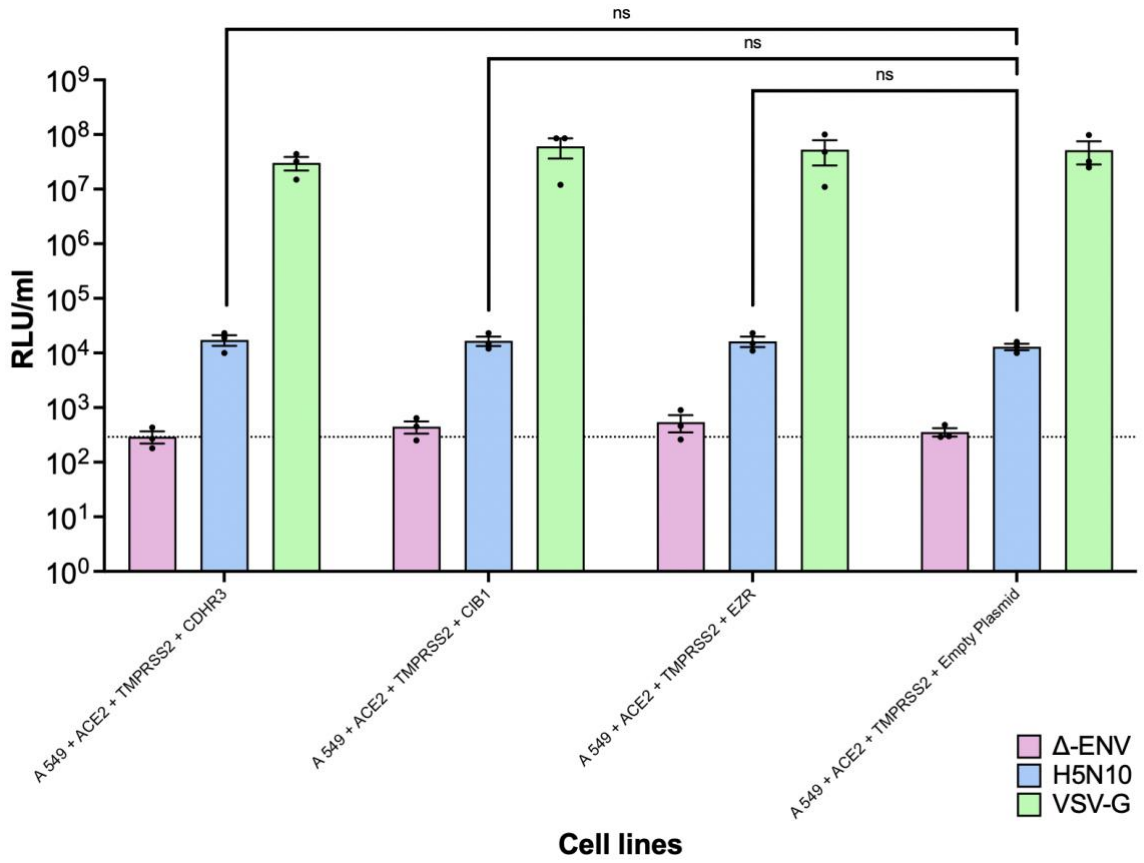



***Figure 28. Effect of CDHR3, CIB1, or EZR overexpression on influenza type A viral isolate H5N10 pseudovirus entry in A549 cells.***

A549 cells continuously expressing ACE2 and TMPRSS2 transduced with H5N10 pseudoviruses following transient transfection with CDHR3, CIB1, EZR or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ No statistically significant differences in luciferase signal were observed between conditions relative to the negative control. Although slightly higher numerical values were detected across all conditions compared to the negative control, these were consistent across transcript overexpression conditions and therefore do not indicate a transcript-specific effect on H5N10 pseudovirus entry. Data are presented as $\log_{10}$ RLU/ml.

## 5.6 Functional evaluation of candidate transcripts in SARS-CoV-2 wild-type and variant pseudovirus entry

SARS-CoV-2 wild-type and variant pseudoviruses were used to evaluate whether candidate transcripts influence viral entry. Overexpression of CDHR3 did not result in a statistically significant increase in pseudovirus entry relative to the negative control across most variants (**Fig. 30**). However, a statistically significant increase in entry was observed for the Alpha variant compared to the corresponding control condition, suggesting a potential variant-specific effect. No significant differences were observed for the remaining variants.

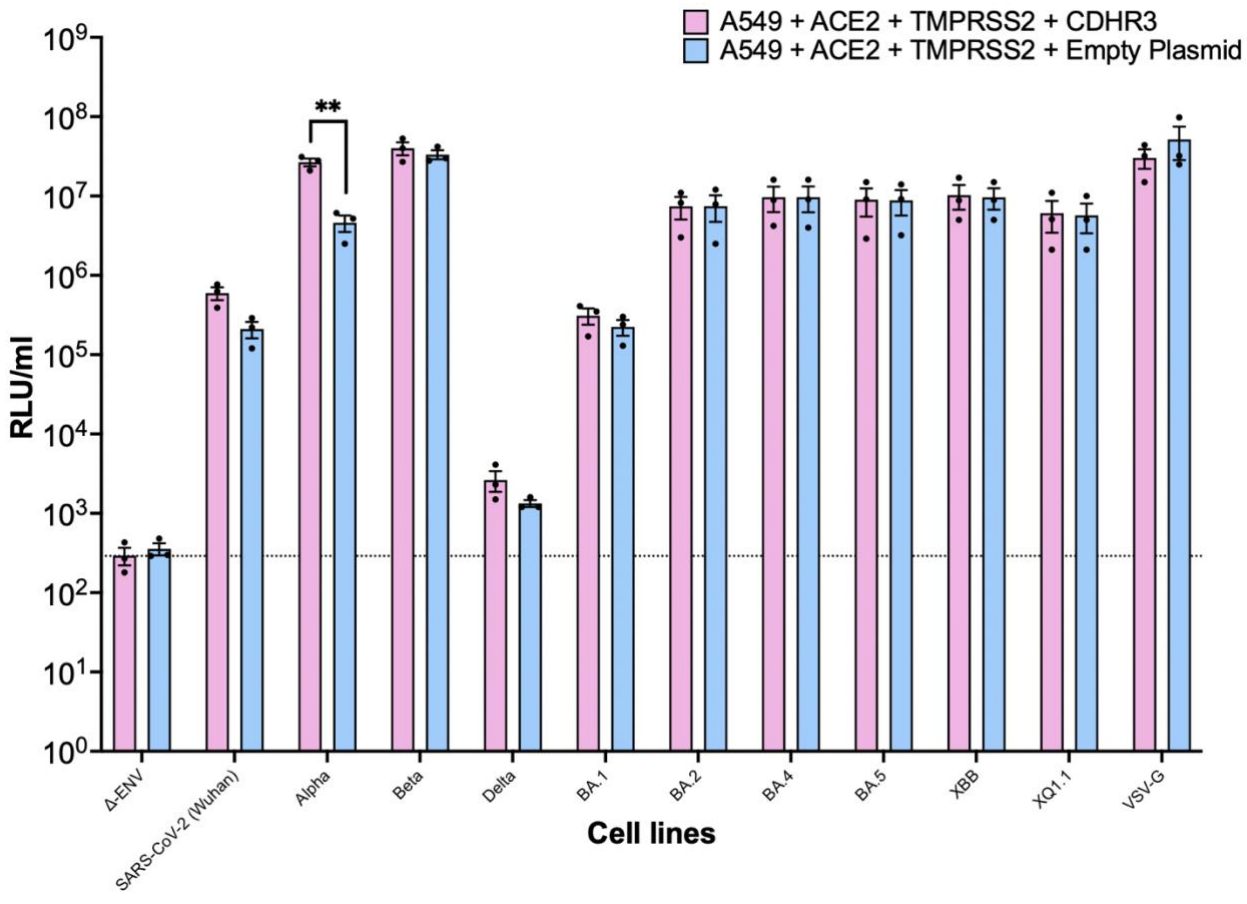



***Figure 29. Effect of CDHR3 overexpression on SARS-CoV-2 wild-type and variant pseudovirus entry in A549 cells.***

A549 cells continuously expressing ACE2 and TMPRSS2 transduced with SARS-CoV-2 wild-type and variant pseudoviruses following transient transfection with CDHR3, or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative

control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. No statistically significant differences were observed relative to the negative control. However, cells overexpressing CDHR3 showed a statistically significant increase in Alpha pseudovirus entry compared to the control condition. Data are presented as $\log_{10}$ RLU/ml.

Overexpression of CIB1 did not result in a statistically significant increase in pseudovirus entry compared to control conditions for any variant tested. Luciferase signals remained comparable across all conditions, indicating no detectable effect of CIB1 on SARS-CoV-2 wild-type and its variant pseudovirus entry.

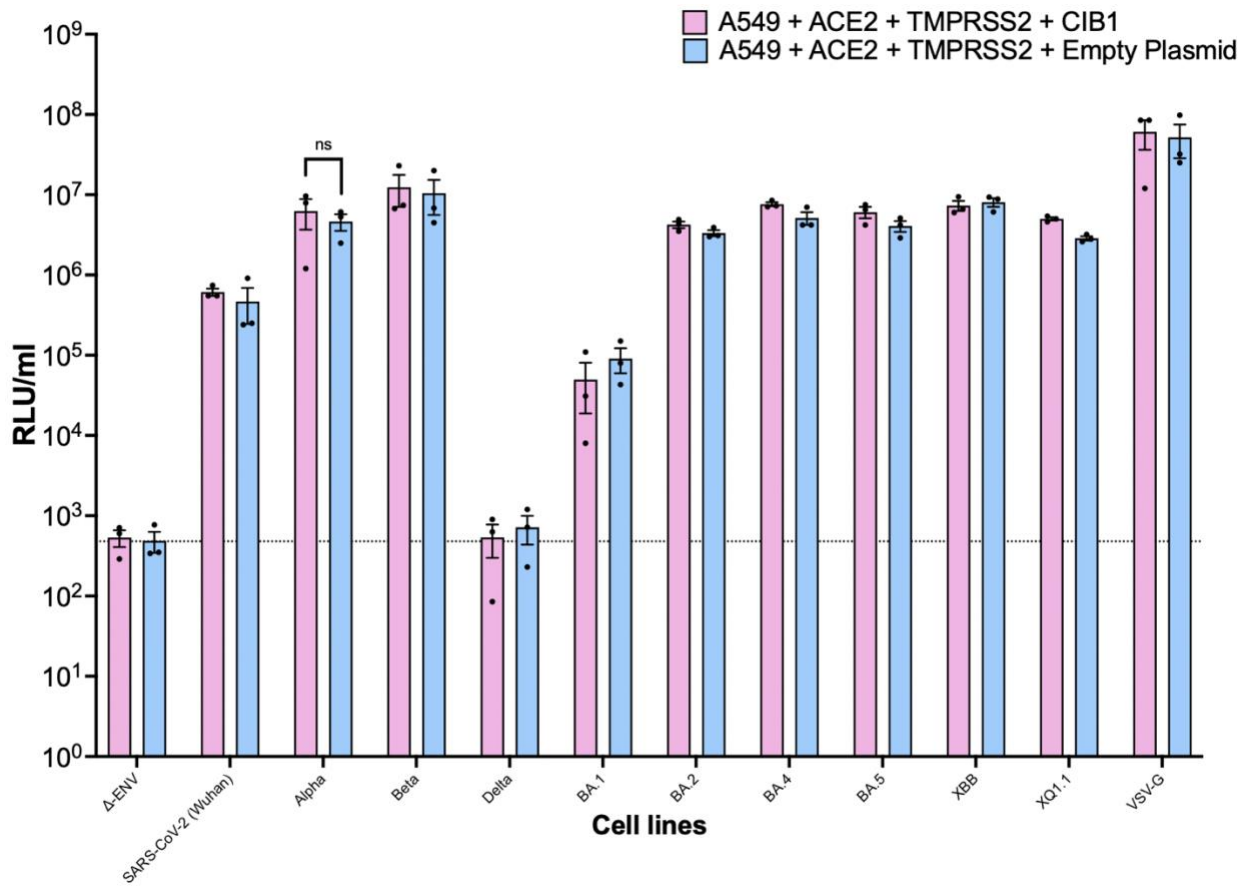



***Figure 30. Effect of CIB1 overexpression on SARS-CoV-2 wild-type and variant pseudovirus entry in A549 cells.***

A549 cells continuously expressing ACE2 and TMPRSS2 transduced with SARS-CoV-2 and variant pseudoviruses following transient transfection with CIB1, or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No statistically significant differences in luciferase signal were observed relative to negative control. Overexpression of CIB1 did not result in a statistically significant increase across the wild-type or variants. Data are presented as $\log_{10}$ RLU/ml.

Similarly, overexpression of EZR did not significantly alter pseudovirus entry across any SARS-CoV-2 variants or the wild-type. Luciferase signals were consistent with control conditions, suggesting that EZR does not play a significant role in viral entry within this system.

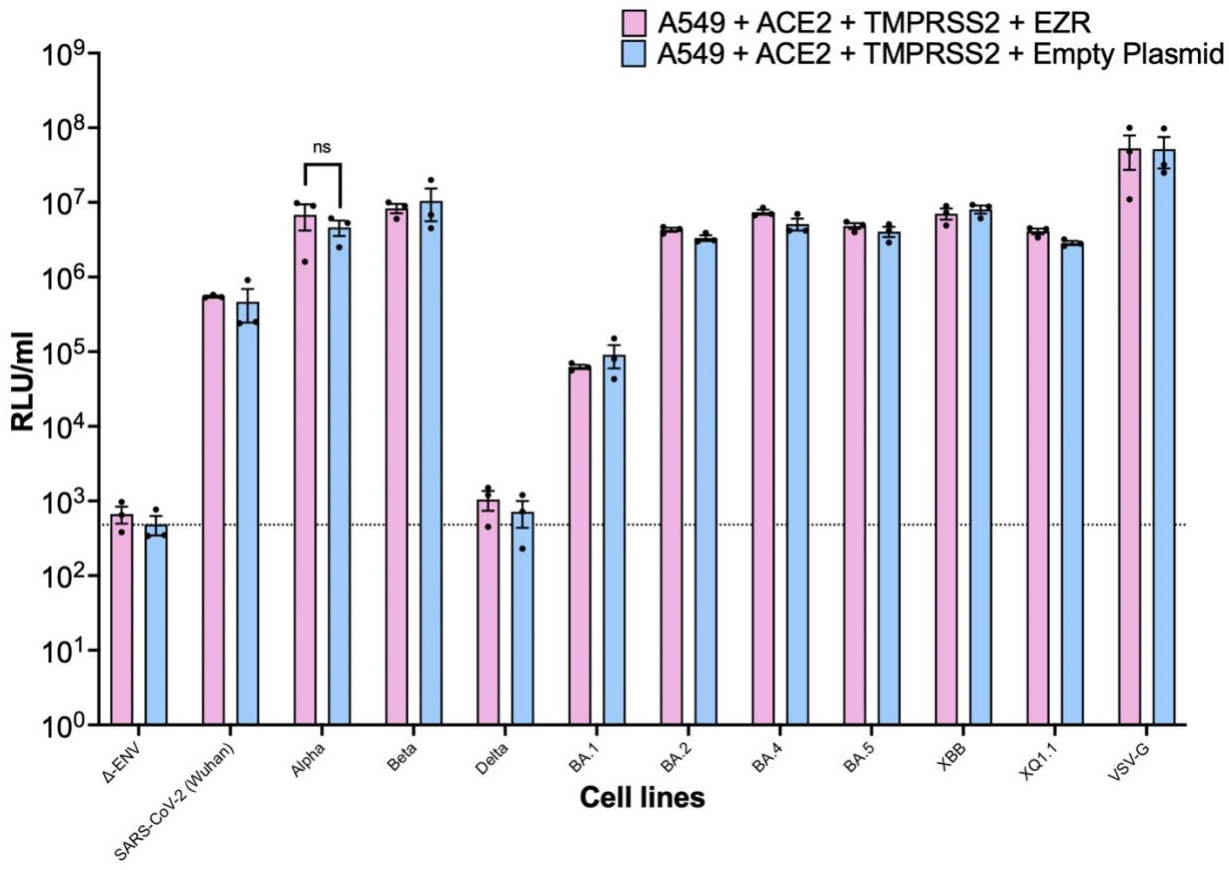



***Figure 31. Effect of EZR overexpression on SARS-CoV-2 wild-type and variant pseudovirus entry in A549 cells.***

A549 cells continuously expressing ACE2 and TMPRSS2 transduced with SARS-CoV-2 and variant pseudoviruses following transient transfection with EZR, or an empty plasmid vector control. VSV-G protein was used as a positive control, and a pseudovirus lacking envelope glycoprotein was included as a negative control. Pseudovirus entry was assessed using a luciferase reporter assay. Data are shown as mean ± SEM from n = 3 independent experiments, each performed with technical duplicates. Statistical significance was determined using two-way ANOVA with Tukey's multiple comparison test. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. No statistically significant differences in luciferase signal were observed relative to negative control. Overexpression of EZR did not result in a statistically significant increase across the wild-type or variants. Data are presented as $\log_{10}$ RLU/ml.

## 5.7 Relative expression of candidate genes in target cell lines using qPCR

Overexpression of CDHR3, CIB1, and EZR in A549 + ACE2 + TMPRSS2 cell lines was confirmed by qPCR. Relative expression levels were determined by the ΔΔCt method, with HPRT as the reference gene and A549 + ACE2 + TMPRSS2 cells with no overexpression of the candidate genes as the control sample (**Fig. 33**). CIB1 expression showed a high $\log_2$ fold change compared to CDHR3 and EZR, while CDHR3 and EZR exhibited lower but significant fold changes. These results confirm a successful upregulation of all three genes in this cell line. Melting curve analysis demonstrated a single peak for each target gene and the reference gene (HPRT), confirming amplification specificity (**Fig. 34**).

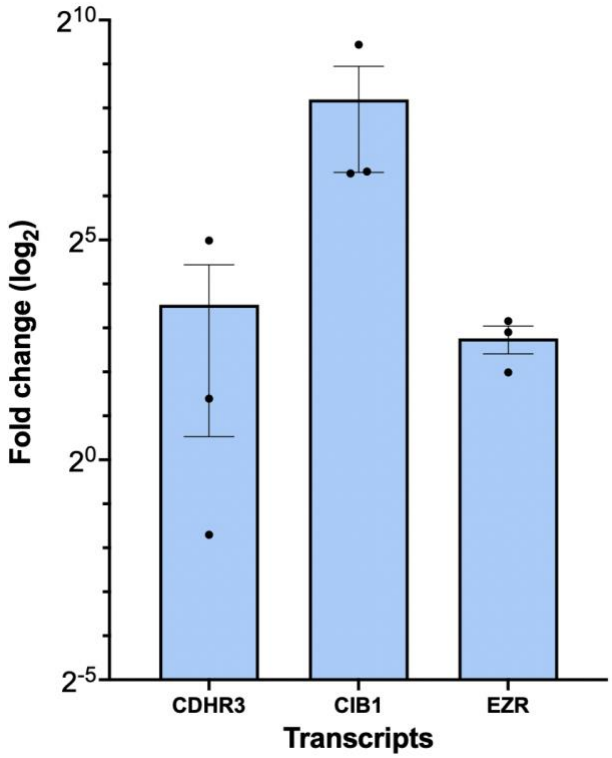



***Figure 32 Relative gene expression of CDHR3, CIB1, and EZR in A549 + ACE2 + TMPRSS2 cell line.*** Expression levels were determined by qPCR, normalised to HPRT, and calculated relative to non-overexpressing control A549 + ACE2 + TMPRSS2 cells using the ΔΔCt method.

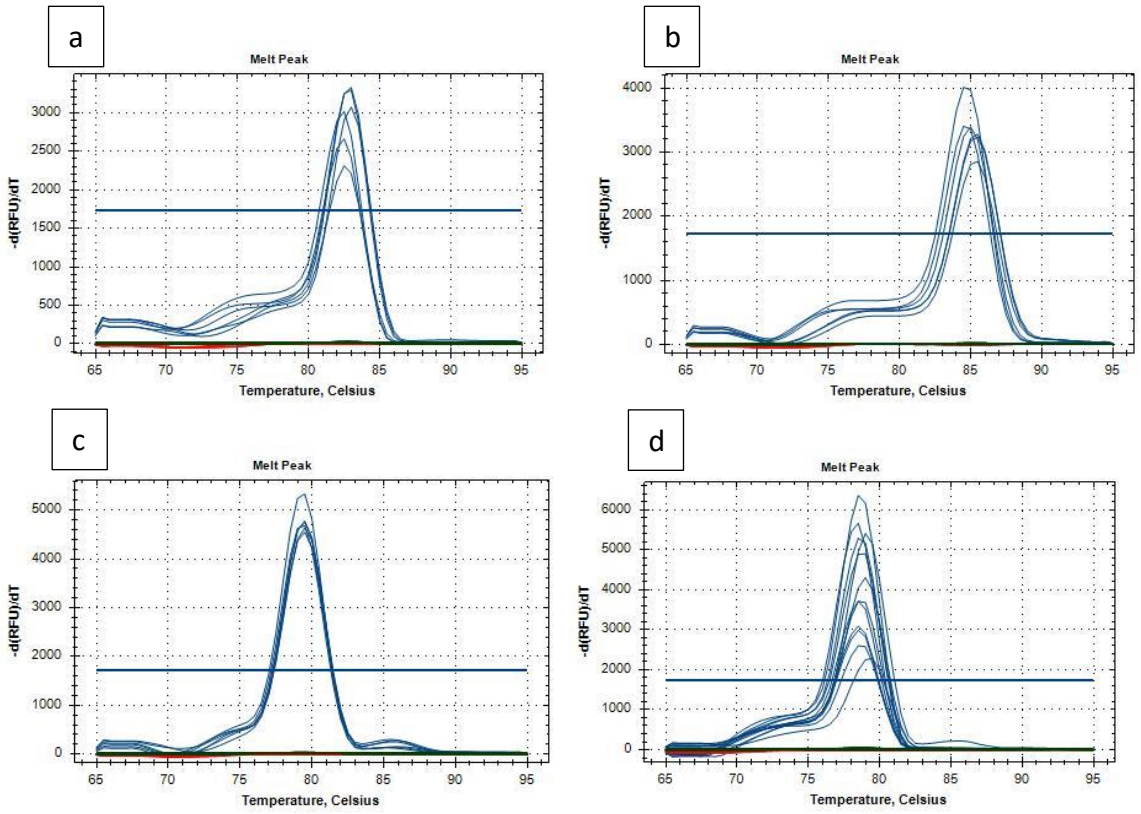



***Figure 33. qPCR melting curve analysis.***

Melting curves for (**a**) CDHR3, (**b**) CIB1, (**c**) EZR, and (**d**) HPRT showing single peaks, indicating specific amplification and absence of primer-dimer formation.

## 5.8 Interpretation of cross-virus entry assays

Pseudovirus entry validation experiments confirmed that SARS-CoV-1, SARS-CoV-2, and influenza A (H5N10) pseudoviruses were capable of efficient entry in ACE2 and TMPRSS2 overexpressing cells, whereas MERS-CoV pseudoviruses exhibited no entry, consistent with the absence of their required host receptors. Within the HKU1 model, no detectable pseudovirus entry was observed across any condition in three independent biological replicates (**Fig. 25 and 26**). A single experiment showed entry in CDHR3 and CDHR3 + TMPRSS2 overexpressing cells; however, this was not reproducible across replicates and is therefore not considered biologically meaningful.

In contrast, SARS-CoV-1 and influenza A pseudovirus entry was consistently observed; however, overexpression of CDHR3, CIB1, or EZR did not result in statistically significant differences compared to control conditions in either of the systems. MERS-CoV pseudoviruses did not exhibit detectable entry under the conditions tested, consistent with the lack of its entry receptor. As such, MERS-CoV served as a negative control, confirming that observed effects in other systems were not due to non-specific enhancement of viral entry.

Analysis of SARS-CoV-2 wild-type and variant pseudoviruses revealed a limited and variant-specific effect. While overexpression of CIB1 and EZR did not significantly alter viral entry, CDHR3 overexpression resulted in a statistically significant increase in entry of the Alpha variant compared to the corresponding control condition. No significant effects were observed for other variants, suggesting that any potential role for CDHR3 in viral entry may be variant-dependent rather than broadly applicable across different coronaviruses and their variants.

Overall, these results demonstrate that transcript enrichment in respiratory epithelial cells does not necessarily translate to a functional role in viral entry when assessed through overexpression alone. Furthermore, the findings highlight the critical influence of receptor

availability and cell line permissiveness in pseudovirus systems, which must be carefully considered when interpreting negative results.

# Chapter 6: Discussion

This thesis aimed to identify host factors underlying respiratory virus tropism for ciliated epithelial cells, with the hypothesis that transcripts enriched in these cells may contribute to viral entry. Using a combination of transcriptomic analysis, CDHR3, CIB1, and EZR were prioritised and tested in the lab using pseudovirus model systems. Overall, this study provides insight into the relationship between ciliated cell transcript enrichment and viral entry. The following sections discuss the technical challenges, candidate selection, and functional findings in the context of current literature.

## 6.1 Challenges in studying HCoV-HKU1 entry

Infectious diseases have a huge impact on public health globally due to their high pathogenicity. Many viruses must therefore be handled in BSL 3 or 4 containment laboratories to carry out research (Li *et al.*, 2018). Alongside this, viruses such as HKU1 are difficult to culture in standard cell lines. Live virus has, however, been successfully used to infected primary cells, showing specific tropism for ciliated cells (Pyrc *et al.*, 2010). Despite this, HKU1 remains incredibly difficult to culture, combined with limited availability of live viral isolates and the lack of established reverse genetic systems to generate recombinant virus from sequence data. To overcome these limitations, pseudovirus systems have become widely used tools for investigating viral attachment and entry, particularly following the emergence of SARS-CoV-2 (Trischitta *et al.*, 2024). Pseudoviruses are viral particles, which are specifically designed to only express surface characteristics of enveloped viruses, without the genomic DNA or RNA, which makes them replication-deficient (Thimmiraju, Kimata and Pollet, 2024). Therefore, it is safe to carry out research in lower BSL laboratories, which are more readily available, allowing more scientists to get involved in infectious disease research (Bentley, Mather and Temperton, 2015).

In this study, pseudoviruses were generated using the HIV-1-based packaging system, incorporating gag-pol, viral spike glycoproteins, and a firefly luciferase reporter to quantify viral entry (Li *et al.*, 2018). This approach enabled the functional investigation of candidate host factors under controlled conditions while maintaining safety and accessibility.

Despite these advantages, significant challenges were encountered when attempting to model entry of HKU1. Despite extensive optimisation efforts, significant pseudovirus entry could not be established. This is consistent with previous reports highlighting the difficulty of propagating and experimentally modelling HKU1 *in vitro*, which has limited detailed characterisation of its entry mechanisms (Pyrc *et al.*, 2010). HKU1 is thought to utilise TMPRSS2 for attachment and may require additional, poorly defined host factors or specific cellular environments that are not readily replicated in standard immortalised cell lines (Saunders *et al.*, 2023).

Multiple strategies were explored, including spike truncation, alternative packaging cell lines, protease supplementation, pseudovirus concentration, and incorporation of additional structural proteins. However, none resulted in detectable HKU1 pseudovirus entry. Additionally, overexpression of TMPRSS2 alone was insufficient to support pseudovirus transduction, suggesting that HKU1 entry may depend on a more complex entry mechanism.

While optimisation approaches covered a vast range of experimental variables, they do not encompass all possible conditions. Additional factors, including the use of more alternative target cell lines, different transfection reagents, or a broader range of host proteases, may influence pseudovirus entry and were not fully assessed in this study.

Overall, these findings highlight both the strengths and limitations of pseudovirus-based approaches. While pseudoviruses offer a powerful and safe platform for studying viral entry, they may not fully capture the complexity of entry processes for all viruses due to not all factors needed being present. In the case of HKU1, the inability to establish a reliable

entry model suggests that it may require specific factors or cellular compositions not recreated in the models used here.

## 6.2 Identification and prioritisation of host factors in ciliated cells

Following transcriptomic filtering and subcellular localisation analysis, a subset of candidate host factors upregulated in respiratory ciliated cells was prioritised (see Chapter 3, Table 5). Additionally, factors that were downregulated in ciliated cells were analysed, and top 13 candidates were selected, although these were not experimentally validated (see Chapter 3, Figure 10). Selection criteria included enrichment in ciliated epithelial cells, and membrane localisation. To choose the final three candidates' evidence from prior literature was gathered, focusing on potential involvement in virus-host interactions.

This approach enabled systematic reduction of a large transcript list to a focused set of biologically relevant candidates. Prioritising membrane-associated proteins was particularly important, as viral entry is mediated by interactions at the cell surface. In addition, incorporating existing literature increased the likelihood of identifying functionally relevant candidates with possible roles in viral entry.

Importantly, this approach focused primarily on upregulated transcripts in ciliated cells, whereas downregulated genes may also contribute to viral entry. Therefore, while this filtering strategy provided a rational framework for candidate selection, it may not fully capture the complexity of host factors influencing respiratory virus tropism.

## 6.3 Functional evaluation of candidate transcripts

Following the selection of candidate transcripts enriched in ciliated epithelial cells, experimental validation was undertaken using pseudovirus-based entry assays to assess

whether CDHR3, CIB1, or EZR influence viral entry when overexpressed in immortalised cell lines under controlled *in vitro* conditions.

The inability to establish a reliable HCoV-HKU1 pseudovirus system limited direct evaluation of candidate involvement in HKU1 entry. However, in a single experimental replicate, an increase in luciferase signal was observed in CHO cells overexpressing CDHR3, with a further increase in cell co-expressing CDHR3 and TMPRSS2, while TMPRSS2 alone had no observable effect. Although this pattern may suggest a potential role for CDHR3 in facilitating HKU1 pseudovirus entry, the lack of reproducibility across independent experiments limits the strength of this observation. Therefore, this finding is interpreted with caution and may reflect experimental variability rather than a consistent biological effect.

To enable broader assessment, additional respiratory virus pseudovirus systems were employed, including multiple coronaviruses, influenza A virus, and SARS-CoV-2 variants, many of which preferentially target ciliated cells in the respiratory epithelium. Across these systems, overexpression of CDHR3, CIB1, and EZR did not result in statistically significant increases in viral entry compared to control cell lines for most viral strains tested. These findings suggest that these candidates do not function as general entry factors or co-factors in the cell models tested, including ACE2 and TMPRSS2 overexpressing immortalised epithelial cells. Although the overexpression was confirmed by qPCR, it may still have been insufficient to fully support their activity. However, it is important to note that CDHR3 is a well-established receptor for HRV-C (Everman *et al.*, 2019), supporting the idea that it may facilitate virus-specific entry mechanisms rather than acting as a generalised entry factor across respiratory viruses. The absence of consistent effects across multiple viruses in this study further supports a context-dependent role.

Importantly, a statistically significant increase in viral entry was observed for the SARS-CoV-2 Alpha variant in CDHR3 overexpressing cells. The Alpha variant has a strong

tropism for ciliated epithelial cells, and is characterised by spike mutations associated with increased transmissibility and altered receptor binding affinity (Davies *et al.*, 2021; Starr *et al.*, 2020; Liu *et al.*, 2021b), which may influence interactions with host cell surface proteins and enhance entry efficiency within these cells. Therefore, suggesting that CDHR3 may be the reason for increased entry for the Alpha variant compared to wild-type or other variants. However, as this effect was not observed for other SARS-CoV-2 variants or respiratory viruses, it is unlikely to represent a generalisable role for CDHR3 and instead suggests a potentially context-dependent interaction requiring further validation.

The absence of a consistent CDHR3-dependent effect across viruses further supports the fact that its established role as a viral receptor is likely virus-specific rather than broadly applicable to respiratory viruses. Similarly, although CIB1 and EZR have been implicated in virus-host interactions in other contexts, their overexpression alone was insufficient to enhance viral entry, suggesting that they are unlikely to directly mediate entry under these conditions tested. However, this does not exclude ciliated cell-specific tropism, as such tropism may also be influenced by the unique pattern of glycans on the cell surface of ciliated airway epithelial cells, including the distribution of sialic acid-containing glycans that are known to contribute to attachment of respiratory viruses such as influenza viruses and some coronaviruses, including HKU1 (Kumlin *et al.*, 2008; Hulswit *et al.*, 2019; Thompson *et al.*, 2006). Instead, the roles of CIB1 and EZR may be associated with downstream stages of infection.

Overall, these results indicate that high transcript enrichment in ciliated epithelial cells does not necessarily mean that a transcript has a direct functional role in viral entry. While CDHR3, CIB1, and EZR are enriched in respiratory ciliated epithelial cells and are supported by prior literature in viral entry involvement, their overexpression did not translate into measurable changes in pseudovirus systems under the conditions tested in this study.

This portrays the complexity of viral entry mechanisms and suggests that additional factors are critical determinants of infection.

## 6.4 Limitations and future work

Several limitations of this study should be acknowledged, which also inform key directions for future work.

First, the transcriptomic analyses relied on publicly available single-cell RNA seq datasets, which introduces variability due to differences in sample collection, sequencing platforms, and data processing pipelines. In addition, the datasets analysed were derived exclusively from adult donors and included a relatively small sample size (n=26), limiting generalisability across age groups, health backgrounds, and ethnicity. This is particularly relevant given the established association between CDHR3 expression, HRV-C infection, and childhood asthma (Bochkov *et al.*, 2015). Future studies incorporating larger, more diverse cohorts, including paediatric samples, would strengthen these findings.

A further limitation lies on the reliance on transcriptomic data to infer protein-level function. Transcript enrichment does not necessarily correlate with protein expression, localisation, or functional activity at the cell surface (Liu, Beyer and Aebersold, 2016). This is particularly important in the context of viral entry, which depends on the presence of functional receptors and co-factors at the plasma membrane. Future work should therefore include complementary approaches such as single-cell proteomics or spatial proteomics to directly quantify protein expression and localisation (Slavov, 2023).

Additionally, the candidate selection strategy depended on annotated subcellular localisation databases, such as SubcellulaRVis, which may bias analysis toward well-characterised proteins. As a result, potentially relevant but less well-characterised proteins

may have been excluded at this step. Integrating multiple localisation resources and unbiased proteomic datasets could mitigate this limitation and expand candidate discovery.

An additional limitation relates to the selection of the HKU1 spike CTD for in silico docking analyses. The CTD was selected because it contains the primary receptor-binding region of the spike protein and was therefore considered the most likely region to interact with host cell factors. However, this assumption may not be true for additional factors. If the factors bind to the same region as the primary receptor, then focusing on the CTD may be appropriate. If the factors interact with distinct regions of the spike protein, such as the NTD or other surface regions, the docking strategy used here would not be useful to identify these interactions. Additionally, the choice of CTD may have biased candidate identification towards proteins predicted to interact with the RBD rather than capturing the full range of potential host interactions. Future studies could address this limitation by performing docking analyses across the entire spike ectodomain and comparing predicted binding sites. Finally, assessing whether the modelling approach can correctly predict interactions with established coronavirus entry factors, such as TMPRSS2, would provide a useful benchmark for evaluating the biological relevance and predictive power of the docking methodology.

Importantly, the assumption that viral entry factors are proteinaceous may not fully be true. Increasing evidence suggests that cell surface glycans, play critical roles in viral attachment and entry, particularly for respiratory viruses such as influenza A virus and coronaviruses (Tortorici *et al.*, 2019). These features are not captured by transcriptomic analyses. Future studies could therefore employ glycan arrays or glycoproteomics to identify specific glycan structures enriched on ciliated cells that may mediate viral interactions.

Limitations are also associated with the experimental systems used for functional validation. While pseudovirus-based assays provide a safe and flexible platform for studying viral entry, they may not fully represent the complexity of authentic virus infection. In

particular, pseudoviruses lack viral replication and may not capture contributions from viral proteins beyond those involved in entry. This limitation was seen in attempts to model HKU1 entry, where extensive optimisation failed to produce effective pseudovirus transduction. These findings suggest that HKU1 entry may depend on additional host factors, co-receptors, or cellular contexts that are not easily reproduced in standard immortalised cell lines (Pyrc *et al.*, 2010).

The pseudovirus system used in this study also requires successful completion of multiple sequential steps, including viral attachment, internalisation, and membrane fusion. Therefore, a negative result does not identify which stage of entry is impaired. Future studies could employ complementary approaches that isolate individual stages of the entry process. For example, fluorescently labelled spike proteins could be used to study attachment independently of internalisation (Tortorici *et al.*, 2019), while binding assays, flow cytometry-based receptor engagement studies, or live-cell imaging could be used to distinguish attachment defects from downstream entry events (Clausen *et al.*, 2020),

A further consideration is that incorporation of HKU1 spike into pseudoviral particles was not directly confirmed. While transfection efficiency was monitored and pseudovirus production protocols followed established methods, failure to demonstrate spike incorporation introduces uncertainty when interpreting the absence of successful HKU1 pseudovirus transduction. If spike incorporation was inefficient or absent, the observed results may reflect a technical limitation of pseudovirus production rather than a true biological requirement for additional host factors. Future studies should therefore verify spike incorporation using approaches such as protein pull-down assays, or mass spectrometry-based analyses. In addition, the use of alternative pseudovirus platforms may improve spike incorporation and functionality, while performing infection assays at temperatures more representative of the URT may better reflect the physiological conditions under which HKU1

naturally spreads. Expanding candidate selection to include lower-abundance transcripts and ciliogenesis-associated genes may also identify additional host factors that contribute to the selective tropism of HKU1 for ciliated airways cells, which may aid entry.

Immortalised epithelial cell lines do not fully reflect the state of airway epithelial cells *in vivo*. Future work could therefore utilise primary human airway epithelial cultures grown at an air-liquid interface, which more closely mimic the structure and function of respiratory epithelium, including ciliation and mucus production (Ravindra *et al.*, 2021). While overexpression studies may be less feasible in these systems, scRNA-seq could be used to characterise primary cells heterogenicity and validate expression patterns.

Technical considerations may also have contributed to variability in experimental outcomes. Although overexpression of candidate transcripts was confirmed at the mRNA level by qPCR, protein expression and cellular localisation were not directly assessed. This is particularly important because successful viral entry depends on the presence of correct proteins at the cell surface. Without confirmation of protein expression and localisation, it is difficult to determine whether negative findings reflect a lack of biological involvement or insufficient surface expression of the candidate proteins. Future work therefore could incorporate protein-level verification using techniques such as Western blotting, flow cytometry, immunofluorescence microscopy, or cell-surface biotinylation assays.

To further validated candidate host factors, complementary functional approaches should be used. Loss-of-function studies, such as siRNA or CRISPR-Cas9 mediated knockdown or knockout, would help determine whether these factors are required for viral entry rather than simply sufficient under overexpression conditions. Additionally, including the use of protease inhibitors and live-cell imaging of fluorescently labelled viral particles, could further reveal viral entry pathways and the role of host cell surface proteins.

Where available, the use of authentic viral isolates would provide a more physiologically relevant assessment of transcript involvement in viral entry. Although access to viruses such as HKU1 is limited, studies using live virus systems could help validate findings from pseudovirus assays and clarify whether candidate transcripts influence entry, replication, or downstream stages of infection.

Finally, further investigation into the potential role of CDHR3 in viral entry, particularly in the context of specific viral variants, may help clarify whether the observations reported here reflect a biologically relevant interaction or experimental variability.

## 6.5 Conclusion

This study combined transcriptomic analysis, *in silico* modelling, and functional pseudovirus assays to investigate host factors that may contribute to respiratory virus tropism for ciliated epithelial cells. A systematic filtering strategy identified candidate transcripts enriched in ciliated cells, leading to the prioritisation of CDHR3, CIB1, and EZR for experimental evaluation.

Functional assessment demonstrated that overexpression of these candidates did not result in consistent enhancement of viral entry across multiple respiratory viruses and SARS-CoV-2 variants. Although a statistically significant increase in entry was observed for the SARS-CoV-2 Alpha variant in CDHR3-overexpressing cells, this effect was not reproduced across other viruses or variants, suggesting a limited or context-dependent role.

The inability to establish a robust HCoV-HKU1 pseudovirus system further highlighted the limitations of current *in vitro* models and suggests that additional host factors or specific cellular environments may be required to support HKU1 entry.

Overall, these findings indicate that transcript enrichment in ciliated epithelial cells alone is insufficient to define functional involvement in viral entry when assessed by

overexpression in simplified cell systems. Instead, viral tropism likely depends on a combination of receptor availability, cellular context, and virus-specific interactions. While CDHR3, CIB1, and EZR remain biologically relevant candidates based on their expression patterns and reported virus associations, their roles may act at stages beyond initial entry and require further investigation.

This study provides a structured framework for the identification, prioritisation and functional evaluation of host factors involved in respiratory virus infection of airway ciliated epithelial cells. The approach used here informs future investigations using more physiologically relevant models and contributes to a broader understanding of the factors shaping viral entry at the respiratory epithelium.

References

Adivitiya, Kaushik, M. S., Chakraborty, S., Veleri, S. and Kateriya, S. (2021) 'Mucociliary respiratory epithelium integrity in molecular defense and susceptibility to pulmonary viral infections', *Biology,* 10(2), pp. 95.

Ahn, J. H., Kim, J., Hong, S. P., Choi, S. Y., Yang, M. J., Ju, Y. S., Kim, Y. T., Kim, H. M., Rahman, M. T. and Chung, M. K. (2021) 'Nasal ciliated cells are primary targets for SARS-CoV-2 replication in the early stage of COVID-19', *The Journal of Clinical Investigation,* 131(13).

Akello, J. O., Kamgang, R., Barbani, M. T., Suter-Riniker, F., Aebi, C., Beuret, C., Paris, D. H., Leib, S. L. and Ramette, A. (2021) 'Genomic analyses of human adenoviruses unravel novel recombinant genotypes associated with severe infections in pediatric patients', *Scientific Reports,* 11(1), pp. 24038.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A. and Wishart, D. S. (2016) 'Heatmapper: web-enabled heat mapping for all', *Nucleic Acids Research,* 44(W1), pp. W147-W153.

Bandyopadhyay, C., Valiya-Veettil, M., Dutta, D., Chakraborty, S. and Chandran, B. (2014) 'CIB1 synergizes with EphrinA2 to regulate Kaposi's sarcoma-associated herpesvirus macropinocytic entry in human microvascular dermal endothelial cells', *PLoS Pathogens,* 10(2), pp. e1003941.

Basnet, S., Bochkov, Y. A., Brockman-Schneider, R. A., Kuipers, I., Aesif, S. W., Jackson, D. J., Lemanske Jr, R. F., Ober, C., Palmenberg, A. C. and Gern, J. E. (2019) 'CDHR3 asthma-risk genotype affects susceptibility of airway epithelium to rhinovirus C infections', *American Journal of Respiratory Cell and Molecular Biology,* 61(4), pp. 450-458.

Battles, M. B. and McLellan, J. S. (2019) 'Respiratory syncytial virus entry and how to block it', *Nature Reviews Microbiology,* 17(4), pp. 233-245.

Belouzard, S., Millet, J. K., Licitra, B. N. and Whittaker, G. R. (2012) 'Mechanisms of coronavirus cell entry mediated by the viral spike protein', *Viruses,* 4(6), pp. 1011-1033.

Bentley, E. M., Mather, S. T. and Temperton, N. J. (2015) 'The use of pseudotypes to study viruses, virus sero-epidemiology and vaccination', *Vaccine,* 33(26), pp. 2955-2962.

Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., Welsch, K., Winkler, M., Schneider, H. and Hofmann-Winkler, H. (2013) 'TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium', *Journal of Virology,* 87(11), pp. 6150-6160.

Bochkov, Y. A., Watters, K., Ashraf, S., Griggs, T. F., Devries, M. K., Jackson, D. J., Palmenberg, A. C. and Gern, J. E. (2015) 'Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication', *Proceedings of the National Academy of Sciences,* 112(17), pp. 5485-5490.

Boncristiani, H., Criado, M. and Arruda, E. (2009) 'Respiratory viruses', *Encyclopedia of Microbiology*, pp. 500.

Bost, P. and Drayman, N. (2024) 'Dissecting viral infections, one cell at a time, by single-cell technologies', *Microbes and Infection,* 26(7), pp. 105268.

Brant, A. C., Tian, W., Majerciak, V., Yang, W. and Zheng, Z.-M. (2021) 'SARS-CoV-2: from its discovery to genome structure, transcription, and replication', *Cell & Bioscience,* 11, pp. 1-17.

Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., Teichmann, S. A., Marioni, J. C. and Stegle, O. (2015) 'Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells', *Nature Biotechnology,* 33(2), pp. 155-160.

Buqaileh, R., Saternos, H., Ley, S., Aranda, A., Forero, K. and AbouAlaiwi, W. A. (2021) 'Can cilia provide an entry gateway for SARS-CoV-2 to human ciliated cells?', *Physiological Genomics,* 53(6), pp. 249-258.

Caì, Y., Postnikova, E. N., Bernbaum, J. G., Yú, S., Mazur, S., Deiuliis, N. M., Radoshitzky, S. R., Lackemeyer, M. G., McCluskey, A. and Robinson, P. J. (2015) 'Simian hemorrhagic fever virus cell entry is dependent on CD163 and uses a clathrin-mediated endocytosis-like pathway', *Journal of Virology,* 89(1), pp. 844-856.

'CALM1', (2025). Available at: https://www.proteinatlas.org/ENSG00000198668-CALM1.

Cambi, A., Koopman, M. and Figdor, C. G. (2005) 'How C-type lectins detect pathogens', *Cellular Microbiology,* 7(4), pp. 481-488.

Carabelli, A. M., Peacock, T. P., Thorne, L. G., Harvey, W. T., Hughes, J., 6, C.-G. U. C. d. S. T. I., Peacock, S. J., Barclay, W. S., de Silva, T. I. and Towers, G. J. (2023) 'SARS-CoV-2 variant biology: immune escape, transmission and fitness', *Nature Reviews Microbiology,* 21(3), pp. 162-177.

Casasnovas, J. M. (2013) 'Virus-receptor interactions and receptor-mediated virus entry into host cells', *Structure and Physics of Viruses: An Integrated Textbook*, pp. 441-466.

'CCDC78', (2025). Available at: https://www.proteinatlas.org/ENSG00000162004-CCDC78.

'CDHR3', (2025). Available at: https://www.proteinatlas.org/ENSG00000128536-CDHR3.

Chan, C. M., Lau, S. K., Woo, P. C., Tse, H., Zheng, B.-J., Chen, L., Huang, J.-D. and Yuen, K.-Y. (2009) 'Identification of major histocompatibility complex class IC molecule as an attachment factor that facilitates coronavirus HKU1 spike-mediated infection', *Journal of Virology,* 83(2), pp. 1026-1035.

Chavda, V. P., Bezbaruah, R., Valu, D., Patel, B., Kumar, A., Prasad, S., Kakoti, B. B., Kaushik, A. and Jesawadawala, M. (2023) 'Adenoviral vector-based vaccine platform for COVID-19: Current status', *Vaccines,* 11(2), pp. 432.

Chen, H.-Y., Huang, C., Tian, L., Huang, X., Zhang, C., Llewellyn, G. N., Rogers, G. L., Andresen, K., O'Gorman, M. R. and Chen, Y.-W. (2021) 'Cytoplasmic tail truncation of SARS-CoV-2 spike protein enhances titer of pseudotyped vectors but masks the effect of the D614G mutation', *Journal of Virology,* 95(22), pp. 10.1128/jvi. 00966-21.

Chen, M. and Zhang, X.-E. (2021) 'Construction and applications of SARS-CoV-2 pseudoviruses: a mini review', *International Journal of Biological Sciences,* 17(6), pp. 1574.

Chen, Y., Ou, X., Li, P., Zan, F., Tan, L. and Qian, Z. (2024) 'Identification of the critical residues of TMPRSS2 for entry and host range of human coronavirus HKU1', *Journal of Virology,* 98(12), pp. e01587-24.

Chu, H., Chan, J. F.-W., Wang, Y., Yuen, T. T.-T., Chai, Y., Hou, Y., Shuai, H., Yang, D., Hu, B. and Huang, X. (2020) 'Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19', *Clinical Infectious Diseases,* 71(6), pp. 1400-1409.

Churchill, M. J., Deeks, S. G., Margolis, D. M., Siliciano, R. F. and Swanstrom, R. (2016) 'HIV reservoirs: what, where and how to target them', *Nature Reviews Microbiology,* 14(1), pp. 55-60.

'CIB1', (2025). Available at: https://www.proteinatlas.org/ENSG00000185043-CIB1.

Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W.-C., Wang, C.-B. and Bernardini, S. (2020) 'The COVID-19 pandemic', *Critical Reviews in Clinical Laboratory Sciences,* 57(6), pp. 365-388.

Clausen, T. M., Sandoval, D. R., Spliid, C. B., Pihl, J., Perrett, H. R., Painter, C. D., Narayanan, A., Majowicz, S. A., Kwong, E. M. and McVicar, R. N. (2020) 'SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2', *Cell,* 183(4), pp. 1043-1057. e15.

Coate, K. C., Cha, J., Shrestha, S., Wang, W., Gonçalves, L. M., Almaça, J., Kapp, M. E., Fasolino, M., Morgan, A. and Dai, C. (2020) 'SARS-CoV-2 cell entry factors ACE2 and TMPRSS2 are expressed in the microvasculature and ducts of human pancreas but are not enriched in β cells', *Cell Metabolism,* 32(6), pp. 1028-1040. e4.

Compans, R. W. and Herrler, G. (2007) 'Virus infection of epithelial cells', *Mucosal Immunology*, pp. 769.

*Coronaviruses* (2024): National Institute of Allergy and Infectious Diseases. Available at: https://www.niaid.nih.gov/diseases-conditions/coronaviruses (Accessed: 3rd January 2025).

Correa, D., Giraldo, D. M., Gallego, S., Taborda, N. A. and Hernandez, J. C. (2023) 'Immunity towards human respiratory syncytial virus', *Acta Virologica,* 67, pp. 11887.

Cossart, P. and Helenius, A. (2014) 'Endocytosis of viruses and bacteria', *Cold Spring Harbor Perspectives in Biology,* 6(8), pp. a016972.

Dai, X., Xu, R. and Li, N. (2024) 'The Interplay between Airway Cilia and Coronavirus Infection, Implications for Prevention and Control of Airway Viral Infections', *Cells,* 13(16), pp. 1353.

Davidson, M. W. (2015) *The Influenza (Flu) Virus*. Molecular Expressions: The Flroida State University. Available at: https://micro.magnet.fsu.edu/cells/viruses/influenzavirus.html (Accessed: 18th of June 2025).

Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., Pearson, C. A., Russell, T. W., Tully, D. C. and Washburne, A. D. (2021) 'Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England', *Science,* 372(6538), pp. eabg3055.

Davis, J. D. and Wypych, T. P. (2021) 'Cellular and functional heterogeneity of the airway epithelium', *Mucosal Immunology,* 14(5), pp. 978-990.

Deprez, M., Zaragosi, L.-E., Truchi, M., Becavin, C., Ruiz García, S., Arguel, M.-J., Plaisant, M., Magnone, V., Lebrigand, K. and Abelanet, S. (2020) 'A single-cell atlas of the human healthy airways', *American Journal of Respiratory and Critical Care Medicine,* 202(12), pp. 1636-1645.

Dimitrov, D. S. (2004) 'Virus entry: molecular mechanisms and biomedical applications', *Nature Reviews Microbiology,* 2(2), pp. 109-122.

Divala, T. (2023) *What is a variant? An expert explains*: Wellcome. Available at: https://wellcome.org/news/what-variant-expert-explains (Accessed: 26th of May 2025).

'DNAAF1', (2025). Available at: https://www.proteinatlas.org/ENSG00000154099-DNAAF1.

Dou, D., Revol, R., Östbye, H., Wang, H. and Daniels, R. (2018) 'Influenza A virus cell entry, replication, virion assembly and movement', *Frontiers in Immunology,* 9, pp. 1581.

Doyle, C. A., Busey, G. W., Iobst, W. H., Kiessling, V., Renken, C., Doppalapudi, H., Stremska, M. E., Manjegowda, M. C., Arish, M. and Wang, W. (2024) 'Endosomal fusion of pH-dependent enveloped viruses requires ion channel TRPM7', *Nature Communications,* 15(1), pp. 8479.

'DYNLL1', (2025). Available at: https://www.proteinatlas.org/ENSG00000088986-DYNLL1.

El-Sayed, A. and Kamel, M. (2021) 'Coronaviruses in humans and animals: the role of bats in viral evolution', *Environmental Science and Pollution Research,* 28(16), pp. 19589-19600.

Escalante, D. E., Wang, A. and Ferguson, D. M. (2021) 'Structural Modeling of the TMPRSS Subfamily of Host Cell Proteases Reveals Potential Binding Sites', *bioRxiv*, pp. 2021.06. 15.448583.

Everman, J. L., Sajuthi, S., Saef, B., Rios, C., Stoner, A. M., Numata, M., Hu, D., Eng, C., Oh, S. and Rodriguez-Santana, J. (2019) 'Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations', *Journal of Allergy and Clinical Immunology,* 144(4), pp. 962-971.

'EZR', (2025). Available at: https://www.proteinatlas.org/ENSG00000092820-EZR.

Farhud, D. D. and Mojahed, N. (2022) 'SARS-COV-2 notable mutations and variants: a review article', *Iranian Journal of Public Health,* 51(7), pp. 1494.

Fehr, A. R. and Perlman, S. (2015) 'Coronaviruses: an overview of their replication and pathogenesis', *Coronaviruses: Methods and Protocols*, pp. 1-23.

Feldman, S. A., Audet, S. and Beeler, J. A. (2000) 'The fusion glycoprotein of human respiratory syncytial virus facilitates virus attachment and infectivity via an interaction with cellular heparan sulfate', *Journal of Virology,* 74(14), pp. 6442-6447.

Feng, F., Zhu, Y., Ma, Y., Wang, Y., Yu, Y., Sun, X., Song, Y., Shao, Z., Huang, X. and Liao, Y. (2023) 'A CRISPR activation screen identifies genes that enhance SARS-CoV-2 infection', *Protein & Cell,* 14(1), pp. 64-68.

Forni, D., Cagliani, R., Pozzoli, U., Mozzi, A., Arrigoni, F., De Gioia, L., Clerici, M. and Sironi, M. (2022) 'Dating the emergence of human endemic coronaviruses', *Viruses,* 14(5), pp. 1095.

Godinho-Santos, A., Hance, A. J., Gonçalves, J. and Mammano, F. (2016) 'CIB1 and CIB2 are HIV-1 helper factors involved in viral entry', *Scientific Reports,* 6(1), pp. 30927.

Goldberg, A. R., Langwig, K. E., Brown, K. L., Marano, J. M., Rai, P., King, K. M., Sharp, A. K., Ceci, A., Kailing, C. D. and Kailing, M. J. (2024) 'Widespread exposure to SARS-CoV-2 in wildlife communities', *Nature Communications,* 15(1), pp. 6210.

Goldstein, S. A. and Weiss, S. R. (2017) 'Origins and pathogenesis of Middle East respiratory syndrome-associated coronavirus: recent advances', *F1000Research,* 6.

Griggs, T. F., Bochkov, Y. A., Basnet, S., Pasic, T. R., Brockman-Schneider, R. A., Palmenberg, A. C. and Gern, J. E. (2017) 'Rhinovirus C targets ciliated airway epithelial cells', *Respiratory Research,* 18, pp. 1-11.

Guarner, J. 2020. Three emerging coronaviruses in two decades: the story of SARS, MERS, and now COVID-19. Oxford University Press US.

Hamre, D. and Procknow, J. J. (1966) 'A new virus isolated from the human respiratory tract', *Proceedings of the Society for Experimental Biology and Medicine,* 121(1), pp. 190-193.

Hardenbrook, N. J. and Zhang, P. (2022) 'A structural view of the SARS-CoV-2 virus and its assembly', *Current Opinion in Virology,* 52, pp. 123-134.

Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., Ludden, C., Reeve, R., Rambaut, A. and Consortium, C.-G. U. (2021) 'SARS-CoV-2 variants, spike mutations and immune escape', *Nature Reviews Microbiology,* 19(7), pp. 409-424.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H. and Nitsche, A. (2020) 'SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor', *Cell,* 181(2), pp. 271-280. e8.

Horton, K., Wing, P. A., Jackson, C. L., McCormick, C. J., Carroll, M. P. and Lucas, J. S. (2025) 'Interplay between respiratory viruses and cilia in the airways', *European Respiratory Review,* 34(175).

'HSP90AA1', (2025). Available at: https://www.proteinatlas.org/ENSG00000080824-HSP90AA1.

Huang, W., Wang, D. and Yao, Y.-F. (2021) 'Understanding the pathogenesis of infectious diseases by single-cell RNA sequencing', *Microbial Cell,* 8(9), pp. 208.

Huang, X., Dong, W., Milewska, A., Golda, A., Qi, Y., Zhu, Q. K., Marasco, W. A., Baric, R. S., Sims, A. C. and Pyrc, K. (2015) 'Human coronavirus HKU1 spike protein uses O-acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme', *Journal of Virology,* 89(14), pp. 7202-7213.

Huang, Y., Yang, C., Xu, X.-f., Xu, W. and Liu, S.-w. (2020) 'Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19', *Acta Pharmacologica Sinica,* 41(9), pp. 1141-1149.

Hulswit, R. J., Lang, Y., Bakkers, M. J., Li, W., Li, Z., Schouten, A., Ophorst, B., van Kuppeveld, F. J., Boons, G.-J. and Bosch, B.-J. (2019) 'Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A', *Proceedings of the National Academy of Sciences,* 116(7), pp. 2681-2690.

Hutchinson, E. C., von Kirchbach, J. C., Gog, J. R. and Digard, P. (2010) 'Genome packaging in influenza A virus', *Journal of General Virology,* 91(2), pp. 313-328.

Hwang, B.-Y. and Schaffer, D. V. (2013) 'Engineering a serum-resistant and thermostable vesicular stomatitis virus G glycoprotein for pseudotyping retroviral and lentiviral vectors', *Gene Therapy,* 20(8), pp. 807-815.

Ison, M. G. and Hayden, R. T. (2016) 'Adenovirus', *Diagnostic Microbiology of the Immunocompromised Host*, pp. 217-232.

Jackson, C. B., Farzan, M., Chen, B. and Choe, H. (2022) 'Mechanisms of SARS-CoV-2 entry into cells', *Nature Reviews Molecular Cell Biology,* 23(1), pp. 3-20.

Jacobs, S. E., Lamson, D. M., St. George, K. and Walsh, T. J. (2013) 'Human rhinoviruses', *Clinical Microbiology Reviews,* 26(1), pp. 135-162.

Javanian, M., Barary, M., Ghebrehewet, S., Koppolu, V., Vasigala, V. and Ebrahimpour, S. (2021) 'A brief review of influenza virus infection', *Journal of Medical Virology,* 93(8), pp. 4638-4646.

Jones, M. S., Harrach, B., Ganac, R. D., Gozum, M. M., dela Cruz, W. P., Riedel, B., Pan, C., Delwart, E. L. and Schnurr, D. P. (2007) 'New adenovirus species found in a patient presenting with gastroenteritis', *Journal of Virology,* 81(11), pp. 5978-5984.

Jovic, D., Liang, X., Zeng, H., Lin, L., Xu, F. and Luo, Y. (2022) 'Single-cell RNA sequencing technologies and applications: A brief overview', *Clinical and Translational Medicine,* 12(3), pp. e694.

Kelly, J. N., Laloli, L., V'kovski, P., Holwerda, M., Portmann, J., Thiel, V. and Dijkman, R. (2022) 'Comprehensive single cell analysis of pandemic influenza A virus infection in the human airways uncovers cell-type specific host transcriptional signatures relevant for disease progression and pathogenesis', *Frontiers in Immunology,* 13, pp. 978824.

Kia'i, N. and Bajaj, T. (2019) 'Histology, respiratory epithelium'.

Kirchdoerfer, R. N., Cottrell, C. A., Wang, N., Pallesen, J., Yassine, H. M., Turner, H. L., Corbett, K. S., Graham, B. S., McLellan, J. S. and Ward, A. B. (2016) 'Pre-fusion structure of a human coronavirus spike protein', *Nature,* 531(7592), pp. 118-121.

Klasse, P. J., Bron, R. and Marsh, M. (1998) 'Mechanisms of enveloped virus entry into animal cells', *Advanced Drug Delivery Reviews,* 34(1), pp. 65-91.

Krammer, F., Palese, P., Shaw, M., García-Sastre, A., Smith, G., Fouchier, R., Peiris, M., Kedzierska, K., Doherty, P. and Treanor, J. (2018) 'Influenza', *Nature Reviews. Disease Primers,* 4(1), pp. 1-21.

Kumlin, U., Olofsson, S., Dimock, K. and Arnberg, N. (2008) 'Sialic acid tissue distribution and influenza virus tropism', *Influenza and Other Respiratory Viruses,* 2(5), pp. 147-154.

Lasaro, M. O. and Ertl, H. C. (2009) 'New insights on adenovirus as vaccine vectors', *Molecular Therapy,* 17(8), pp. 1333-1339.

Lee, W.-M., Lemanske Jr, R. F., Evans, M. D., Vang, F., Pappas, T., Gangnon, R., Jackson, D. J. and Gern, J. E. (2012) 'Human rhinovirus species and season of infection determine illness severity', *American Journal of Respiratory and Critical Care Medicine,* 186(9), pp. 886-891.

Leung, H. S., Li, O. T., Chan, R. W., Chan, M. C., Nicholls, J. M. and Poon, L. L. (2012) 'Entry of influenza A Virus with a α2, 6-linked sialic acid binding preference requires host fibronectin', *Journal of Virology,* 86(19), pp. 10704-10713.

Li, C.-X., Noreen, S., Zhang, L.-X., Saeed, M., Wu, P.-F., Ijaz, M., Dai, D.-F., Maqbool, I., Madni, A. and Akram, F. (2022) 'A critical analysis of SARS-CoV-2 (COVID-19) complexities, emerging variants, and therapeutic interventions and vaccination strategies', *Biomedicine & Pharmacotherapy,* 146, pp. 112550.

Li, Q., Liu, Q., Huang, W., Li, X. and Wang, Y. (2018) 'Current status on the development of pseudoviruses for enveloped viruses', *Reviews in Medical Virology,* 28(1), pp. e1963.

Li, Q., Shah, T., Wang, B., Qu, L., Wang, R., Hou, Y., Baloch, Z. and Xia, X. (2023) 'Cross-species transmission, evolution and zoonotic potential of coronaviruses', *Frontiers in Cellular and Infection Microbiology,* 12, pp. 1081370.

Liu, D. X., Liang, J. Q. and Fung, T. S. (2021) 'Human coronavirus-229E,-OC43,-NL63, and-HKU1 (Coronaviridae)', *Encyclopedia of Virology*, pp. 428.

Liu, J., Li, Y., Liu, Q., Yao, Q., Wang, X., Zhang, H., Chen, R., Ren, L., Min, J. and Deng, F. (2021a) 'SARS-CoV-2 cell tropism and multiorgan infection', *Cell Discovery,* 7(1), pp. 17.

Liu, Y., Beyer, A. and Aebersold, R. (2016) 'On the dependency of cellular protein levels on mRNA abundance', *Cell,* 165(3), pp. 535-550.

Liu, Y., Liu, J., Plante, K. S., Plante, J. A., Xie, X., Zhang, X., Ku, Z., An, Z., Scharton, D. and Schindewolf, C. (2021b) 'The N501Y spike substitution enhances SARS-CoV-2 transmission', *BioRxiv*.

Liu, Y.-G., Chen, Y., Wang, X., Zhao, P., Zhu, Y. and Qi, Z. (2020) 'Ezrin is essential for the entry of Japanese encephalitis virus into the human brain microvascular endothelial cells', *Emerging Microbes & Infections,* 9(1), pp. 1330-1341.

Lubinski, B. and Whittaker, G. R. (2023) 'The SARS-CoV-2 furin cleavage site: natural selection or smoking gun?', *The Lancet Microbe,* 4(8), pp. e570.

Luo, M. (2011) 'Influenza virus entry', *Viral Molecular Machines*, pp. 201-221.

Lynch, J. P., Fishbein, M. and Echavarria, M. 'Adenovirus'. *Seminars in Respiratory and Critical Care Medicine*: © Thieme Medical Publishers, 494-511.

Mackay, I. M. and Arden, K. E. (2015) 'MERS coronavirus: diagnostics, epidemiology and transmission', *Virology Journal,* 12, pp. 1-21.

Man, W. H., de Steenhuijsen Piters, W. A. and Bogaert, D. (2017) 'The microbiota of the respiratory tract: gatekeeper to respiratory health', *Nature Reviews Microbiology,* 15(5), pp. 259-270.

Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A. and Klenk, H.-D. (2004) 'Human and avian influenza viruses target different cell types in cultures of human airway epithelium', *Proceedings of the National Academy of Sciences,* 101(13), pp. 4620-4624.

McIntosh, K., Dees, J. H., Becker, W. B., Kapikian, A. Z. and Chanock, R. M. (1967) 'Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease', *Proceedings of the National Academy of Sciences,* 57(4), pp. 933-940.

Meier, O. and Greber, U. F. (2004) 'Adenovirus endocytosis', *The Journal of Gene Medicine: A cross-disciplinary journal for research on the science of gene transfer and its clinical applications,* 6(S1), pp. S152-S163.

Milewska, A., Falkowski, K., Kulczycka, M., Bielecka, E., Naskalska, A., Mak, P., Lesner, A., Ochman, M., Urlik, M. and Diamandis, E. (2020) 'Kallikrein 13 serves as a priming protease during infection by the human coronavirus HKU1', *Science Signaling,* 13(659), pp. eaba9902.

Millet, J. K., Kien, F., Cheung, C.-Y., Siu, Y.-L., Chan, W.-L., Li, H., Leung, H.-L., Jaume, M., Bruzzone, R. and Malik Peiris, J. S. (2012) 'Ezrin interacts with the SARS coronavirus Spike protein and restrains infection at the entry stage', *PLoS One,* 7(11), pp. e49566.

Millet, J. K., Tang, T., Nathan, L., Jaimes, J. A., Hsu, H.-L., Daniel, S. and Whittaker, G. R. (2019) 'Production of pseudotyped particles to study highly pathogenic coronaviruses in a biosafety level 2 setting', *Journal of Visualized Experiments: JoVE,* (145).

Millet, J. K. and Whittaker, G. R. (2016) 'Murine leukemia virus (MLV)-based coronavirus spike-pseudotyped particle production and infection', *Bio-protocol,* 6(23), pp. e2035-e2035.

Monto, A. S., Cowling, B. J. and Peiris, J. M. (2014) 'Coronaviruses', *Viral Infections of Humans: Epidemiology and Control*, pp. 199-223.

Morris, G. M. and Lim-Wilby, M. (2008) 'Molecular docking', *Molecular Modeling of Proteins*: Springer, pp. 365-382.

'MS4A8', (2025). Available at: https://www.proteinatlas.org/ENSG00000166959-MS4A8.

NCBI (2025) *TMPRSS2*. National Library of Medicine: National Center for Biotechnology Information. Available at: https://www.ncbi.nlm.nih.gov/gene/7113 (Accessed: 23rd January 2025).

Neerukonda, S. N., Vassell, R., Lusvarghi, S., Liu, S., Akue, A., Kukuruga, M., Wang, T. T., Weiss, C. D. and Wang, W. (2025) 'Characterization of spike S1/S2 processing and entry pathways of lentiviral pseudoviruses bearing seasonal human coronaviruses NL63, 229E, and HKU1 spikes', *Microbiology Spectrum*, pp. e02808-24.

Nuwarda, R. F., Alharbi, A. A. and Kayser, V. (2021) 'An overview of influenza viruses and vaccines', *Vaccines,* 9(9), pp. 1032.

Okuda, K. and Gentzsch, M. 2024. Pulmonary Ionocytes: What Are They Transporting and Which Way? : American Thoracic Society.

Osterhaus, A. D. (2008) 'New respiratory viruses of humans', *The Pediatric Infectious Disease Journal,* 27(10), pp. S71-S74.

Pace, E., Di Vincenzo, S., Ferraro, M., Lanata, L. and Scaglione, F. (2024) 'Role of airway epithelium in viral respiratory infections: Can carbocysteine prevent or mitigate them?', *Immunology,* 172(3), pp. 329-342.

Papi, A., Ison, M. G., Langley, J. M., Lee, D.-G., Leroux-Roels, I., Martinon-Torres, F., Schwarz, T. F., van Zyl-Smit, R. N., Campora, L. and Dezutter, N. (2023) 'Respiratory syncytial virus prefusion F protein vaccine in older adults', *New England Journal of Medicine,* 388(7), pp. 595-608.

Paranjapye, A., Leir, S.-H., Huang, F., Kerschner, J. L. and Harris, A. (2022) 'Cell function and identity revealed by comparative scRNA-seq analysis in human nasal, bronchial and epididymis epithelia', *European Journal of Cell Biology,* 101(3), pp. 151231.

Parry, R. H., Mostafavi, H., McCallum, G., Sng, J. D., Joensuu, M., Short, K. R., Slonchak, A. and Khromykh, A. A. (2025) 'Omicron evolution drives increased ciliated cell tropism and dysfunction in nasal epithelia', *bioRxiv*, pp. 2025.11. 19.689397.

Prelich, G. (2012) 'Gene overexpression: uses, mechanisms, and interpretation', *Genetics,* 190(3), pp. 841-854.

Prichard, K. L., O'Brien, N. S., Murcia, S. R., Baker, J. R. and McCluskey, A. (2022) 'Role of clathrin and dynamin in clathrin mediated endocytosis/synaptic vesicle recycling and implications in neurological diseases', *Frontiers in Cellular Neuroscience,* 15, pp. 754110.

'Protein Data Bank: the single global archive for 3D macromolecular structure data', (2019) *Nucleic Acids Research,* 47(D1), pp. D520-D528.

'PSENEN', (2025). Available at: https://www.proteinatlas.org/ENSG00000205155-PSENEN.

Pyrc, K., Sims, A. C., Dijkman, R., Jebbink, M., Long, C., Deming, D., Donaldson, E., Vabret, A., Baric, R. and van der Hoek, L. (2010) 'Culturing the unculturable: human coronavirus HKU1 infects, replicates, and produces progeny virions in human ciliated airway epithelial cell cultures', *Journal of Virology,* 84(21), pp. 11255-11263.

Qi, J., Zhou, Y., Hua, J., Zhang, L., Bian, J., Liu, B., Zhao, Z. and Jin, S. (2021) 'The scRNA-seq expression profiling of the receptor ACE2 and the cellular protease TMPRSS2 reveals human organs susceptible to SARS-CoV-2 infection', *International Journal of Environmental Research and Public Health,* 18(1), pp. 284.

Ravindra, N. G., Alfajaro, M. M., Gasque, V., Huston, N. C., Wan, H., Szigeti-Buck, K., Yasumoto, Y., Greaney, A. M., Habet, V. and Chow, R. D. (2021) 'Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes', *PLoS Biology,* 19(3), pp. e3001143.

Saunders, N., Fernandez, I., Planchais, C., Michel, V., Rajah, M. M., Baquero Salazar, E., Postal, J., Porrot, F., Guivel-Benhassine, F., Blanc, C., Chauveau-Le Friec, G., Martin, A., Grzelak, L., Oktavia, R. M., Meola, A., Ahouzi, O., Hoover-Watson, H., Prot, M., Delaune, D., ... and Schwartz, O. (2023) 'TMPRSS2 is a functional receptor for human coronavirus HKU1', *Nature,* 624(7990), pp. 207-214.

Schelhaas, M. (2010) 'Come in and take your coat off–how host cells provide endocytosis for virus entry', *Cellular Microbiology,* 12(10), pp. 1378-1388.

Schneider, C., O'Leary, C. E. and Locksley, R. M. (2019) 'Regulation of immune responses by tuft cells', *Nature Reviews Immunology,* 19(9), pp. 584-593.

Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A. and Li, F. (2020a) 'Cell entry mechanisms of SARS-CoV-2', *Proceedings of the National Academy of Sciences,* 117(21), pp. 11727-11734.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A. and Li, F. (2020b) 'Structural basis of receptor recognition by SARS-CoV-2', *Nature,* 581(7807), pp. 221-224.

Shi, T., McAllister, D. A., O'Brien, K. L., Simoes, E. A., Madhi, S. A., Gessner, B. D., Polack, F. P., Balsells, E., Acacio, S. and Aguayo, C. (2017) 'Global, regional, and national disease burden

estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study', *The Lancet,* 390(10098), pp. 946-958.

Shrestha, L. B., Foster, C., Rawlinson, W., Tedla, N. and Bull, R. A. (2022) 'Evolution of the SARS-CoV-2 omicron variants BA. 1 to BA. 5: implications for immune escape and transmission', *Reviews in Medical Virology,* 32(5), pp. e2381.

Singh, M., Bansal, V. and Feschotte, C. (2020) 'A single-cell RNA expression map of human coronavirus entry factors', *Cell Reports,* 32(12).

Slavov, N. (2023) 'Single-cell proteomics: quantifying post-transcriptional regulation during development with mass-spectrometry', *Development,* 150(13), pp. dev201492.

Starr, T. N., Greaney, A. J., Hilton, S. K., Ellis, D., Crawford, K. H., Dingens, A. S., Navarro, M. J., Bowen, J. E., Tortorici, M. A. and Walls, A. C. (2020) 'Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding', *Cell,* 182(5), pp. 1295-1310. e20.

Sungnak, W., Huang, N., Bécavin, C., Berg, M. and Network, H. L. B. (2020) 'SARS-CoV-2 entry genes are most highly expressed in nasal goblet and ciliated cells within human airways', *ArXiv*, pp. arXiv: 2003.06122 v1.

Taubenberger, J. K. and Morens, D. M. (2008) 'The pathology of influenza virus infections', *Annu. Rev. Pathol. Mech. Dis.,* 3(1), pp. 499-522.

Temperton, N. J., Hoschler, K., Major, D., Nicolson, C., Manvell, R., Hien, V. M., Ha, D. Q., De Jong, M., Zambon, M. and Takeuchi, Y. (2007) 'A sensitive retroviral pseudotype assay for influenza H5N1-neutralizing antibodies', *Influenza and Other Respiratory Viruses,* 1(3), pp. 105-112.

Temperton, N. J., Wright, E. and Scott, S. D. (2015) 'Retroviral pseudotypes—from scientific tools to clinical utility', *Encyclopedia Life Sci,* 10(9780470015902), pp. a0021549.

Thevenon, A., Mazza, A.-M., Husbands, J. and Sharples, F. (2015) *Potential risks and benefits of gain-of-function research: summary of a workshop.* National Academies Press.

Thimmiraju, S. R., Kimata, J. T. and Pollet, J. (2024) 'Pseudoviruses, a safer toolbox for vaccine development against enveloped viruses', *Expert Review of Vaccines,* 23(1), pp. 174-185.

Thompson, A. J. and Paulson, J. C. (2021) 'Adaptation of influenza viruses to human airway receptors', *Journal of Biological Chemistry,* 296, pp. 100017.

Thompson, C. I., Barclay, W. S., Zambon, M. C. and Pickles, R. J. (2006) 'Infection of human airway epithelium by human and avian strains of influenza a virus', *Journal of Virology,* 80(16), pp. 8060-8068.

Tjoonk, N. (2023) *What is a UMAP plot?* Single Cell Discoveries. Available at: https://www.scdiscoveries.com/blog/knowledge/what-is-a-umap-plot/#:~:text=Uniform%20manifold%20approximation%20and%20projection%20(UMAP)%20is,cells%20together%20and%20different%20cells%20further%20apart. (Accessed: 24th of December 2025).

'TMC5', (2025). Available at: https://www.proteinatlas.org/ENSG00000103534-TMC5.

Tortorici, M. A., Walls, A. C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., Boons, G.-J., Bosch, B.-J., Rey, F. A. and de Groot, R. J. (2019) 'Structural basis for human coronavirus attachment to sialic acid receptors', *Nature Structural & Molecular Biology,* 26(6), pp. 481-489.

Trischitta, P., Tamburello, M. P., Venuti, A. and Pennisi, R. (2024) 'Pseudovirus-based systems for screening natural antiviral agents: A comprehensive review', *International Journal of Molecular Sciences,* 25(10), pp. 5188.

'TSPAN1', (2025). Available at: https://www.proteinatlas.org/ENSG00000117472-TSPAN1.

Van Der Hoek, L., Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J., Wolthers, K. C., Wertheim-van Dillen, P. M., Kaandorp, J., Spaargaren, J. and Berkhout, B. (2004) 'Identification of a new human coronavirus', *Nature Medicine,* 10(4), pp. 368-373.

van Riel, D., den Bakker, M. A., Leijten, L. M., Chutinimitkul, S., Munster, V. J., de Wit, E., Rimmelzwaan, G. F., Fouchier, R. A., Osterhaus, A. D. and Kuiken, T. (2010) 'Seasonal and pandemic human influenza viruses attach better to human upper respiratory tract epithelium than avian influenza viruses', *The American Journal of Pathology,* 176(4), pp. 1614-1618.

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T. and Veesler, D. (2020) 'Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein', *Cell,* 181(2), pp. 281-292. e6.

Wang, H., Liu, X., Zhang, X., Zhao, Z., Lu, Y., Pu, D., Zhang, Z., Chen, J., Wang, Y. and Li, M. (2024a) 'TMPRSS2 and glycan receptors synergistically facilitate coronavirus entry', *Cell,* 187(16), pp. 4261-4271. e17.

Wang, W., Guan, J., Ren, M., Li, Z., Ji, W., Chen, R., Xu, Y. and Zhang, S. (2024b) 'High-resolution crystal structure of human coronavirus HKU1 receptor binding domain bound to TMPRSS2 receptor', *hLife,* 2(12), pp. 653-657.

Author (2025) *TrailmakerTM*. Available at: http://app.trailmaker.parsebiosciences.com.

Watson, J., Smith, M., Francavilla, C. and Schwartz, J.-M. (2022) 'SubcellulaRVis: a web-based tool to simplify and visualise subcellular compartment enrichment', *Nucleic Acids Research,* 50(W1), pp. W718-W725.

Welliver, R. C. (2003) 'Respiratory syncytial virus and other respiratory viruses', *The Pediatric Infectious Disease Journal,* 22(2), pp. S6-S12.

Whitsett, J. A. and Alenghat, T. (2015) 'Respiratory epithelial cells orchestrate pulmonary innate immunity', *Nature Immunology,* 16(1), pp. 27-35.

WHO (2025) *Influenza (seasonal)*: World Health Organisation. Available at: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) (Accessed: 2025 24th of June).

Woo, P. C. (2023) *Family: Coronaviridae*: ICTV. Available at: https://ictv.global/report/chapter/coronaviridae/coronaviridae (Accessed: 3rd January 2025).

Woo, P. C., Lau, S. K., Chu, C.-m., Chan, K.-h., Tsoi, H.-w., Huang, Y., Wong, B. H., Poon, R. W., Cai, J. J. and Luk, W.-k. (2005) 'Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia', *Journal of Virology,* 79(2), pp. 884-895.

Wu, C.-T., Lidsky, P. V., Xiao, Y., Cheng, R., Lee, I. T., Nakayama, T., Jiang, S., He, W., Demeter, J. and Knight, M. G. (2023) 'SARS-CoV-2 replication in airway epithelia requires motile cilia and microvillar reprogramming', *Cell,* 186(1), pp. 112-130. e20.

Xiang, Q., Li, L., Wu, J., Tian, M. and Fu, Y. (2022) 'Application of pseudovirus system in the development of vaccine, antiviral-drugs, and neutralizing antibodies', *Microbiological Research,* 258, pp. 126993.

Yamauchi, Y. and Helenius, A. (2013) 'Virus entry at a glance', *Journal of Cell Science,* 126(6), pp. 1289-1295.

Yan, Y., Tao, H., He, J. and Huang, S.-Y. (2020) 'The HDOCK server for integrated protein–protein docking', *Nature Protocols,* 15(5), pp. 1829-1852.

Yang, Y.-L., Lu, Z.-Z., Xie, C., Zhong, L.-Y., Sun, C. and Zeng, M.-S. (2024) 'Unraveling Coronavirus Cell Invasion: The Role of Glycan Receptors in Human Coronavirus-HKU1 Cell Entry', *Infectious Diseases & Immunity*, pp. 10.1097.

Ye, Z.-W., Yuan, S., Yuen, K.-S., Fung, S.-Y., Chan, C.-P. and Jin, D.-Y. (2020) 'Zoonotic origins of human coronaviruses', *International Journal of Biological Sciences,* 16(10), pp. 1686.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B. and Huang, C.-L. (2020) 'A pneumonia outbreak associated with a new coronavirus of probable bat origin', *Nature,* 579(7798), pp. 270-273.

Zhu, Z., Lian, X., Su, X., Wu, W., Marraro, G. A. and Zeng, Y. (2020) '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, pp. 1-14.

'ZMYND10', (2025). Available at: https://www.proteinatlas.org/ENSG00000004838-ZMYND10.

# Appendix

**Table 1. 170 highly enriched transcripts in ciliated cells, localised at the plasma membrane or cytosolic facing membranes.**

| |
|---|
| ABCA13 |
| ABHD2 |
| ACTB |
| ACTG1 |
| AFDN |
| AKAP9 |
| ALCAM |
| ALDH3A1 |
| ALDH3B1 |
| ALOX15 |
| ANXA1 |
| ANXA2 |
| ANXA4 |
| APP |
| AQP3 |
| ARF4 |
| ARHGAP18 |
| ATP1B1 |
| ATP6V1D |
| B2M |
| BASP1 |
| CALM1 |
| CALM2 |
| CAPN2 |
| CATSPERD |
| C6 |
| CCN2 |
| CCDC78 |
| CD9 |
| CD24 |
| CD59 |
| CD63 |
| CD74 |
| CD164 |
| CDHR3 |
| CDHR4 |

| |
|---|
| CFAP65 |
| CFAP126 |
| CFAP157 |
| CFLAR |
| CIB1 |
| CITA |
| CLDN3 |
| CLDN4 |
| CLDN7 |
| CLDN8 |
| CLIC1 |
| CLTA |
| COPRS |
| CP |
| CST6 |
| CYSTM1 |
| CTNND1 |
| CTTN |
| DNAAF1 |
| DNAAF4 |
| DYNAA4 |
| DYNC2H1 |
| DYNLL1 |
| DSP |
| ECRG4 |
| ENAH |
| EPCAM |
| EZR |
| FILIP1 |
| FOLR1 |
| FRMPD2 |
| FTO |
| FXYD3 |
| GLUL |
| GPR162 |
| GSN |
| GSTP1 |
| HINT1 |
| HLA-DPA1 |
| HLA-DRA |
| HLA-DRB1 |
| HSPA1A |

| |
|---|
| HSPA5 |
| HSPA8 |
| HSPB1 |
| HSP90AA1 |
| HSP90AB1 |
| HSP90B1 |
| IGFBP2 |
| JUN |
| KIF21A |
| KRT8 |
| KRT10 |
| KRT18 |
| KRT19 |
| KTN1 |
| LGALS3 |
| LMO7 |
| LRP11 |
| LRTOMT |
| LYPD2 |
| MAL2 |
| MAPK10 |
| MORF4L2 |
| MS4A8 |
| MUC1 |
| MUC4 |
| MUC15 |
| MUC16 |
| MUC20 |
| NELFE |
| NUCB2 |
| OCLN |
| OMG |
| OSBPL6 |
| OSCP1 |
| PDLIM4 |
| PERP |
| PIGR |
| PLPP2 |
| PRKAR1A |
| PROM1 |
| PRSS12 |
| PSCA |

| |
|---|
| PSENEN |
| RAC1 |
| RHOB |
| RIPOR2 |
| RRAD |
| RND3 |
| S100A6 |
| S100A10 |
| S100P |
| SAMHD1 |
| SDC4 |
| SLC27A2 |
| SLC3A2 |
| SLC34A2 |
| SLC44A4 |
| SLC7A2 |
| SOD1 |
| SPA17 |
| SPINT2 |
| SPEF1 |
| SPTAN1 |
| SPTBN1 |
| SRI |
| STOM |
| STOML3 |
| SYNE1 |
| SYNE2 |
| SYAP1 |
| SYT8 |
| TACSTD2 |
| TMBIM6 |
| TMC5 |
| TMEM59 |
| TMEM67 |
| TMEM123 |
| TMEM231 |
| TSPAN1 |
| TSPAN3 |
| TSPAN6 |
| TSPAN19 |
| TUSC3 |
| UBAC1 |

| |
|---|
| UBB |
| UBC |
| VAMP8 |
| VIM |
| VSTM2L |
| YWHAE |
| ZMYND10 |

**Table 2. Transcripts from each donor, which were excluded during localisation using SubcellulaRVis.**

| Donor | Excluded transcripts |
|---|---|
| 1 | AC007906.2, AC008771.1, AL357093.2, ANKRD66, ANKUB1, ARL 3.00, ARMC4, C11orf88, C1orf194, C20orf85, C20orf96, C6orf118, C9orf135, CALML4, CASC1, CCDC153, CCDC17, CCDC173, CCDC190, CCDC30, CRNDE, CTXN1, DTHD1, ECT2L, EFCAB1, EFCAB10, ERICH3, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, GIHCG, KIAA1211L, LDLRAD1, LRRC10B, LRRC46, LRRIQ1, MAP3K19, MLF 1.00, MORN5, MT-ND3, NEK5, PLEKHS1, RIIAD1, SNTN, TCTEX1D2, TEX9,WDR54, WDR60, WDR63, WDR66, WDR78, WDR86-AS1, ZBBX |
| 2 | AC007906.2, AC013264.1, AL121956.6, AL357093.2, ANKRD66, ANKUB1, ARL 3.00, ARMC4, C11orf88, C12orf75, C17orf97, C1orf194, C20orf85, C20orf96, C6orf118, C9orf135, CALML4, CASC1, CCDC153, CCDC17, CCDC173, CCDC190, CCDC30, CRNDE, CTXN1, DTHD1, ECT2L, EFCAB1, EFCAB10, ERICH3, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, GIHCG, KIAA1211L, LDLRAD1, LRRC10B, LRRC46, LRRC6, LRRIQ1, MAATS1, MAP3K19, MLF 1.00, MORN5, NEK5, PLEKHS1, RIIAD1, SAMD15, SNTN, |

| | |
|---|---|
| | SRGAP3-AS2, TCTEX1D1, TCTEX1D2, TEX9, TMEM232, UBXN10, WDR54, WDR60, WDR63, WDR66, WDR78, WDR86-AS1, ZBBX |
| 3 | AC007906.2, AL357093.2, C11orf88, C12orf75, C1orf194, C20orf85, C9orf135, CCDC153, CCDC17, EFCAB1, ERICH3, FAM174A, FAM183A, FAM216B, FAM229B, FAM92B, LDLRAD1, LRRC10B, LRRIQ1, MLF 1.00, MORN5, NEK5, SNTN, TCTEX1D2, WDR66 |
| 4 | AC013264.2, ANKUB1, ARL 3.00, ARMC4, ATPIF1, BCYRN1, C11orf70, C11orf88, C12orf75, C14orf142, C1orf194, C20orf85, C21orf58, C21orf59, C22orf15, C6orf118, C9orf135, C9orf9, CAPS - ENSG00000105519, CARS, CCDC153, CCDC17, CCDC190, CCDC74A, CH17-340M24.3, CRNDE, CTGF, CTXN1, DTHD1, EFCAB1, EFCAB10, ERICH3, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, H2AFJ, HAGHL, HN1L, KLHL6, LRRC10B, LRRC46, LRRIQ1, MAP3K19, MLF 1.00, MORN5, MT-ND2, NEK5, PLEKHS1, PRR15, PRR29, RP11-295M3.4, RP11-356K23.1, SNTN, SRGAP3-AS2, TCTEX1D2, TCTEX1D4, TMEM45B, TTC21A, TTC25, UBXN10, WDR34, WDR38, WDR54, WDR66, WDR78, ZBBX |
| 5 | AC007906.2, AL357093.2, C11orf88, C12orf75, C1orf194, C20orf85, C9orf135, CCDC153, CCDC17, EFCAB1, ERICH3, FAM183A, FAM216B, FAM229B, FAM92B, LDLRAD1, LRRIQ1, MLF 1.00, MORN5, NEK5, SNTN, TCTEX1D2 |
| 6 | AC013264.2, ADGB, ANKRD66, ANKUB1, ATPIF1, BCYRN1, C11orf70, C11orf88, C12orf55, C12orf75, C1orf110, C1orf173, C1orf192, C1orf194, C20orf85, C21orf58, C21orf59, C9orf117, C9orf135, CARS, CASC1, SRGAP3-AS2, CCDC11, CCDC114, CCDC153, CCDC17, CCDC173, CCDC176, CCDC19, CCDC30, CCDC42B, CCDC74A, C4orf22, CRIP1 - ENSG00000213145, CTGF, C5orf42, DTHD1, DYX1C1, EFCAB1, FAM154B, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, WDR34, KIAA1377, LINC01171, ARMC4, |

| | |
|---|---|
| | LRRC10B, LRRC46, LRRC48, ARL 3.00, LRRIQ1, MAATS1, MAP3K19, MLF 1.00, MORN5, PCDP1, RABL5, RP11-356K23.1, TEX9, SNTN, HAGHL, HN1L, TCTEX1D2, TCTEX1D4, C17orf72, TTC18, TTC40, UBXN10, WDR16, WDR38, WDR52, WDR54, WDR60, WDR66, WDR78, WDR96, ZBBX |
| 7 | AC013264.2, ANKRD66, ANKUB1, ARL 3.00, ATPIF1, C11orf70, C11orf88, C12orf75, C14orf142, C1orf110, C1orf173, C1orf192, C1orf194, C1orf222, C20orf85, C21orf58, C21orf59, C6orf118, C9orf117, C9orf135, CARS, CASC1, CCDC104, CCDC11, CCDC114, CCDC153, CCDC17, CCDC173, CCDC176, CCDC19, CCDC30, CCDC42B, CCDC74A, CRIP1 - ENSG00000213145, CTGF, DTHD1, DYX1C1, EFCAB1, FAM154B, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, HAGHL, HN1L, HRASLS2, KIAA1377, LINC01171, LRRC10B, LRRC46, LRRC48, LRRIQ1, MAATS1 ,MAP3K19, MLF 1.00, MORN5, PCDP1, PLEKHS1, RABL5, RP11-356K23.1, RP11-620J15.3, SNTN, TCTEX1D2, TCTEX1D4, TTC18, TTC21A, TTC40, UBXN10, WDR16, WDR34, WDR38, WDR52, WDR54, WDR60, WDR66, WDR78, WDR96, ZBBX |
| 8 | ANKRD66, ANKUB1, ARL 3.00, C10orf107, C11orf70, C11orf88, C12orf75, C14orf142, C1orf173, C1orf192, C1orf194, C20orf85, C21orf59, C2orf40, C9orf117, C9orf135, CCDC153, CCDC17, CCDC176, CCDC19, CCDC42B, CCDC74A, CRNDE, CTXN1, DYX1C1, ECT2L, EFCAB1, FAM154B, FAM166B, FAM174A, FAM183A, FAM216B, FAM229B, FAM92B, LRRC10B, LRRC46, LRRC48, LRRIQ1, MAP3K19, MLF 1.00, MORN5, MT-ATP6, MT-CO1, MT-CO2, MT-CO3, MT-CYB, MT-ND1, MT-ND2, MT-ND3, MT-ND4, MT-ND5, RABL5, RP11-356K23.1, SNTN, SRGAP3-AS2, TCTEX1D2, TCTEX1D4, TMEM66, UBXN10, WDR38, WDR54, WDR78, WDR96 |
| 9 | FHAD1, UBXN10, ATPIF1, CCDC30,F AM183A, TCTEX1D4, CCDC17, WDR78, C1orf173, C1orf194, CCDC19, C1orf192, WDR54, PCDP1, CCDC74A, MAP3K19, |

| | |
|---|---|
| | CCDC173, AC013264.2, SNTN, WDR52, MAATS1, ANKUB1, MLF 1.00, ZBBX, TCTEX1D2, DTHD1, C5orf42, FAM229B, CTGF, C6orf118, RABL5, ORAI2, WDR60, EFCAB1, C9orf135, MORN5, WDR38, C9orf117, TTC18, ARL 3.00, WDR96, PLEKHS1, TTC40, CARS, LRRC10B, RARRES3, LINC01171, KIAA1377, C11orf70, C11orf88, CCDC153, CASC1, LRRIQ1, C12orf75, CCDC42B, WDR66, FAM216B, CCDC176, RP11-356K23.1, CRIP1 - ENSG00000213145, DYX1C1, FAM154B, FAM92B, WDR16, LRRC48, LRRC46, CCDC11, C20orf85, CTXN1, SEPW1, CCDC114, FAM227A, C21orf59, C21orf58 |
| 10 | ANKRD66, ANKUB1, ARL 3.00, ATPIF1, C10orf107, C11orf70, C11orf74, C11orf88, C12orf75, C14orf142, C1orf173, C1orf192, C1orf194, C20orf85, C21orf59, C9orf135, CCDC104, CCDC153, CCDC17, CCDC176, CCDC19, CCDC42B, CCDC74A, CRNDE, CTXN1, DYX1C1, EFCAB1, FAM154B, FAM166B, FAM174A, FAM183A, FAM216B, FAM229B, FAM92B, LRRC10B, LRRC46, LRRIQ1, MAP3K19, MLF 1.00, MORN5, MT-ND3, ORAI2, RABL5, RP11-356K23.1, SEPW1, SHFM1, SNTN, SRGAP3-AS2, TCTEX1D2, TCTEX1D4, TMEM66, UBXN10, WDR38, WDR54, WDR63, WDR66, WDR78 |
| 11 | AC013264.2, ANKRD66, ANKUB1, ARL 3.00, ARMC4, ATPIF1, C11orf70, C11orf88, C14orf142, C17orf72, C1orf173, C1orf192, C1orf194, C20orf85, C21orf58, C21orf59, C22orf15, C6orf118, C9orf117, C9orf135, C9orf9, CARS, CASC1, CCDC11, CCDC114, CCDC153, CCDC17, CCDC173, CCDC176, CCDC19, CCDC30, CCDC42B, CRIP1 - ENSG00000213145, CTXN1, DTHD1, DYX1C1, EFCAB1, EFCAB10, FAM154B, FAM166B, FAM183A, FAM216B, FAM227A, FAM229B, FAM65B, FAM92B, FHAD1, KIAA1377, LINC01171, LRRC10B, LRRC46, LRRC48, LRRC6, LRRIQ1, MAATS1, MAP3K19, MLF 1.00, MORN5, MT-ATP6, MT-CYB, MT-ND2, MT-ND3, PCDP1, RABL5, RP11- |

| | |
|---|---|
| | 356K23.1, SNTN, SRGAP3-AS2, TCTEX1D2, TCTEX1D4, TMEM232, TTC18, UBXN10, WDR16, WDR34, WDR38, WDR52, WDR54, WDR60, WDR63, WDR66, WDR78, WDR96, ZBBX |
| 12 | AC013264.2, ANKRD66, ANKUB1, ARL 3.00, ATPIF1, C11orf70, C11orf88, C12orf75, C1orf110, C1orf173, C1orf192, C1orf194, C20orf85, C21orf59, C5orf42, C6orf118, C9orf117, C9orf135, CARS, CASC1, CCDC11, CCDC114, CCDC153, CCDC17, CCDC173, CCDC176, CCDC19, CCDC30, CCDC42B, CCDC74A, CRIP1 - ENSG00000213145, CTGF, DTHD1, DYX1C1, EFCAB1, FAM154B, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, FHAD1, GPR110, KIAA1377, LINC01171, LRRC10B, LRRC46, LRRC48, LRRIQ1, MAP3K19, MLF 1.00, MORN5, RABL5, RARRES3, RP11-356K23.1, SEPW1, SNTN, TCTEX1D2, TCTEX1D4, TTC18, TTC40, UBXN10, WDR38, WDR52, WDR54, WDR60, WDR66, WDR78, WDR96, ZBBX |
| 13 | ARL 3.00, C11orf88, C12orf75, C1orf194, C20orf85, C9orf135, CALML4, CCDC153, CCDC17, CCDC74A, CTXN1, DTHD1, EFCAB1, EFCAB10, ERICH3, FAM183A, FAM216B, FAM229B, FAM92B, GIHCG, LDLRAD1, LOC101927057, LOC101927596, LOC101928817, LOC105371267, LRRC10B, LRRC46, LRRIQ1, MAP3K19, MLF 1.00, MORN5, RIIAD1, SNTN, TCTEX1D2, TMEM232, WDR38, WDR54, WDR78, WDR86-AS1 |
| 14 | C11orf88, C12orf75, C1orf194, C20orf85, C21orf58, CALML4, CCDC153, CCDC17, CCDC173, DTHD1, EFCAB1, ERICH3, FAM174A, FAM183A, FAM216B, FAM227A, FAM229B, FAM92B, HAGHL, LDLRAD1, LOC101928817, LOC105371267, LOC107986833, LRRC10B, LRRC46, LRRIQ1, MAP3K19, MLF 1.00, MORN5, NEK5, PLEKHS1, SNTN, TCTEX1D2, TCTEX1D4, TMEM232, WDR54, WDR78, ZBBX |

| | |
|---|---|
| 15 | C11orf88, C1orf194, C20orf85, CCDC153, CCDC17, CCDC190, DTHD1, EFCAB1, ERICH3, FAM183A, FAM227A, FAM229B, FAM92B, HAGHL, LDLRAD1, LOC101927057, LOC101928817, LOC105371267, LOC107986833, LRRC10B, LRRC46, LRRIQ1, MLF 1.00, MORN5, SEC14L4, SNTN, TCTEX1D2, TCTEX1D4, TMEM232, WDR54 |
| 16 | C11orf88, C1orf194, C20orf85, C9orf135, CCDC153, CCDC17, CCDC190, DTHD1, EFCAB1, ERICH3, FAM183A, FAM227A, FAM229B, FAM92B, HAGHL, LDLRAD1, LOC101927057, LOC101928817, LOC105371267, LOC107986833, LRRC46, LRRIQ1, MORN5, NEK5, SNTN, TCTEX1D2, TMEM232, WDR54, ZBBX |
| 17 | ATPIF1, C11orf88, C20orf85, CAPS - ENSG00000105519, FAM183A, FAM229B, MORN5, MT-CYB |
| 18 | AC013264.2, ANKRD66, ARL 3.00, ATPIF1, C10orf107, C11orf70, C11orf74, C11orf88, C12orf75, C14orf142, C1orf194, C20orf85, C21orf58, C21orf59, C22orf15, C2orf40, C6orf118, C9orf135, C9orf9, CAPS - ENSG00000105519, CCDC153, CCDC17, CCDC74A, CH17-340M24.3, CRNDE, EFCAB1, EFCAB10, ERICH3, FAM166B, FAM174A, FAM183A, FAM216B, FAM229B, FAM92B, HIST1H1C, LDLRAD1, LRRC46, LRRIQ1, MLF 1.00, MORN5, RIIAD1, RP11-295M3.4, RP11-356K23.1, RP11-620J15.3, RP4-666F24.3, SEPW1, SNTN, TCTEX1D2, TCTEX1D4, WBSCR27, WDR38, WDR54, WRB |
| 19 | AC013264.2, ATPIF1, C11orf88, C12orf75, C1orf194, C20orf85, C APS - ENSG00000105519, CRNDE, EFCAB1, FAM183A, FAM216B, FAM229B, FAM92B, LRRIQ1, MLF 1.00, SNTN |
| 20 | ATPIF1, FAM183A, C1orf194, C2orf40, AC013264.2, SNTN, MLF 1.00, TCTEX1D2, FAM229B, MORN5, C11orf88, LRRIQ1, C12orf75, FAM216B, RP11- |

| | |
|---|---|
| | 356K23.1, RP11-295M3.4, CRNDE, FAM92B, LRRC46, C20orf85, CAPS - ENSG00000105519, SEPW1 |
| 21 | FAM183A, CCDC17, C1orf194, C2orf40, AC013264.1, SNTN, TCTEX1D2, C9orf135, MORN5, CARS, C11orf88, CCDC153, LRRIQ1, AL357093.2, FAM92B, LRRC46, C20orf85, MT-ND5 |
| 22 | FHAD1, CCDC30, FAM183A, CCDC17, LDLRAD1, TCTEX1D1, WDR78, ERICH3, C1orf194, RIIAD1, WDR54, C2orf40, CCDC74A, AC013264.1, MDH1B, SNTN, MAATS1, ANKUB1, ZBBX, WDR49, TCTEX1D2, DTHD1, FAM229B, ADGB, C6orf118, WDR60, AC244090.1, EFCAB1, FAM166B, C9orf135, MORN5, WDR38, C11orf88, CCDC153, LRRIQ1, C12orf75, WDR66, FAM216B, NEK5, HAGHL, CCDC189, AC007906.2, FAM92B, LRRC46, PRR29, C20orf85, CCDC114, FAM227A, C21orf58, MT-ND3, MT-ND4L, MT-ND5 |
| 23 | FAM183A, C1orf194, MLF 1.00, FAM174A, SSSCA1, LRRIQ1, C12orf75, CCPG1, FAM92B, C20orf85 |
| 24 | FHAD1, UBXN10, ATPIF1, CCDC30, FAM183A, TCTEX1D4, LDLRAD1, C1orf87, WDR78, C1orf173, WDR63, C1orf194, RP11-263K19.4, CCDC19, C1orf192, H3F3A, CCDC104, WDR54, C2orf40, PCDP1, CCDC74A, CCDC173, AC013264.2, DALRD3, SNTN, KIAA1407, MAATS1, MLF 1.00, ZBBX, TCTEX1D2, DTHD1, C4orf22, FAM174A, TMEM232, TMEM173, FAM65B, HIST1H1C, ANKRD66, FAM229B, ADGB, C6orf118, C7orf57, GBAS, RABL5, EFCAB10, WDR60, TMEM66, EFCAB1, LRRC6, FAM166B, C9orf135, MORN5, WDR38, C9orf117, C9orf171, C9orf9, ARMC4, C10orf107, TTC18, C10orf95, ARL 3.00, WDR96, TTC40, CARS, C11orf16, LINC01171, KIAA1377, C11orf70, C11orf88, CCDC153, CASC1, RP11-620J15.3, LRRIQ1, |

| | |
|---|---|
| | C12orf55, C12orf75, CCDC42B, FAM216B, CCDC176, RP11-356K23.1, C14orf142, RP11-1008C21.1, DYX1C1, FAM154B, CRNDE, FAM92B, C17orf97, TMEM107, WDR16, LRRC48, LRRC46, C17orf72, CCDC11, C20orf96, C20orf85, PPDPF, C20orf201, PPAP2C, SEPW1, C22orf15, WRB, C21orf58 |
| 25 | FAM183A, CCDC17, LDLRAD1, WDR78, ERICH3, C1orf194, WDR54, AC013264.1, SNTN, MLF 1.00, TCTEX1D2, FAM229B, CTGF, EFCAB1, C9orf135, MORN5, WDR38, CARS, C11orf74, C11orf88, ARL 3.00, LRRIQ1, C12orf75, FAM216B, AL357093.2, AC007906.2, CRNDE, FAM92B, C20orf85, CTXN1 |
| 26 | FAM183A, CCDC17, ERICH3, C1orf194, SNTN, TCTEX1D2, EFCAB1, CARS, LRRIQ1, C12orf75, AL357093.2, AC108134.4, AC007906.2, CCDC114, C20orf85 |